# The Structural Basis for the Regulation of ClC Chloride Channels and Transporters by Cytoplasmic Domains

Dissertation

zur Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät der

Universität Zürich

von

**Sandra Marković**

aus
Serbien

Promotionskomitee

Prof. Dr. Raimund Dutzler (Vorsitz)
Prof. Dr. Ben Schuler
Prof. Dr. Blanche Schwappach

Zürich, 2010

# Acknowledgements

The work described in this PhD thesis was carried out in the group of Prof. Raimund Dutzler, at the Department of Biochemistry, University of Zurich.

I am particularly thankful to Raimund for accepting me as a PhD student in his laboratory and for giving me the opportunity to work on this exciting project. I would like to express my gratitude for his ideas, support, attention and motivation during my PhD thesis.

I am grateful to Prof. Blanche Schwappach, Prof. Ben Schuler and Prof. Markus Grütter for forming my PhD thesis committee, for scientific discussion during the committee meetings and critical reviewing of this thesis.

Many thanks to all the members of the Dutzler lab for their help, support and discussions concerning my project. For the nice atmosphere in the L38 lab and a lot of scientific input I am very thankful to my lab mates Stefan and Eric.

My special thanks go to the members of the ClC team, Sara and Sebastian, with whom I have shared challenging goals in our mutual work on this interesting protein family over the last four years. I would like to thank Sebastian for being always ready to answer many questions I had during my initial experiments in the lab and the work on the ClC-Ka domain. I would like to thank Sara for our collaborative work on the eukaryotic ClC expression, the recent work on the prokaryotic homolog Rm1ClC and for growing beautiful crystals of the Rm1ClC cytoplasmic domain. Additionally, I want to thank Thamar and Marina for their contribution to the ClC project during their Master theses.

Thanks to all the colleagues at the institute who have helped with scientific advices or by providing the necessary infrastructure; especially to Beat Blattmann and his coworkers for running the excellent crystallization facility.

Without the constant encouragement and support coming from my family during my whole educational time I would not have been able to reach this final step and this is why they deserved a big THANK YOU! At the end, I want to thank my husband Marcus for bringing more life into my PhD life and for showing me how beautiful Zurich and Switzerland can be!

# Contents

# Summary

ClC chloride channels and transporters catalyze the selective flow of chloride ions across the membranes of prokaryotic and eukaryotic cells. In eukaryotes, ClC proteins play important roles in the stabilization of membrane potential, transepithelial ion transport and vesicular acidification. The unique feature of the ClC family is that it contains chloride selective channels and secondary active chloride/proton antiporters. Despite this functional diversity, the molecular architecture of the family is conserved consisting of a transmembrane transport domain and a cytoplasmic regulatory domain.

Cytoplasmic domains in eukaryotic ClC proteins have been shown to be crucial for protein function and to be involved in the functional regulation of chloride transport. They have been associated with processes like the regulation of channel gating in response to ligand binding. Since the interaction between the cytoplasmic domains within the homodimeric protein is thought to be crucial for the structural basis of regulation, the aim of my PhD thesis was to reveal the oligomeric assembly of the cytoplasmic domains in eukaryotic ClC proteins. The structure of the cytoplasmic domain of the human channel ClC-Ka determined in this work has been valuable in this respect. The unusual mode of dimerization found in the crystals of the ClC-Ka domain was confirmed to be preserved in solution by combining mutagenesis, cross-linking and analytical ultracentrifugation. The complementary experiments on the ClC-0 cytoplasmic domain have shown that this domain exhibits similar structural organization and that the interaction is likely preserved in the context of the full-length ClC-0 protein. In this way, the previously ambiguous oligomeric structure of the cytoplasmic components in eukaryotic ClC proteins has been clarified.

In order to relate the conformational changes in the cytoplasmic domain with the transmembrane pore, the structure of a full length ClC protein is necessary. However, the current structural information on the ClC family exists only for the transmembrane and cytoplasmic domains in isolation. Therefore, an additional aim of my PhD thesis was to establish the overexpression and purification of a ClC protein containing cytoplasmic domain that would serve as a model systems for structural and functional studies. For that purpose, a set of prokaryotic ClC homologs containing cytoplasmic domains was identified and investigated with respect to their expression in *E. coli*, purification and crystallization behaviour. After extensive examination a ClC homolog from the bacterium *Ralstonia metallidurans*, named Rm1ClC, was selected as the only suitable for crystallization screening. Although broad crystallization attempts were performed for the full-length

Rm1ClC, no crystals have been obtained possibly because of the flexible connections between the cytoplasmic domains and the transmembrane part. For that reason, the transmembrane and the cytoplasmic domain of Rm1ClC were expressed and crystallized in isolation. While the poorly diffracting crystals of the Rm1ClC transmembrane domain have not allowed its detailed structure determination, the structure of the Rm1ClC cytoplasmic domain was determined at 1.6 Å resolution providing a detailed insight into the structural organization of the prokaryotic ClC cytoplasmic domain. Despite the typical CBS fold of this domain, the dimeric arrangement found to be conserved in the eukaryotic ClC proteins was not observed. The Rm1ClC domain was shown to be in a monomeric state in solution indicating at a potential different arrangement between the cytoplasmic domains in prokaryotic ClC proteins.

In summary, the studies presented in this work provide the structural and biochemical information on the oligomeric assembly of the cytoplasmic components in eukaryotic and prokaryotic ClC proteins. This information will serve as the structural basis for future investigations that may reveal the mechanisms of regulation by the cytoplasmic domains in ClC chloride channels and transporters.

# Zusammenfassung

Die Familie der ClC Chloridkanäle und -transporter ist für die spezifische Translokation von Chloridionen über die Membran pro- und eukaryotischer Zellen verantwortlich. In Eukaryoten spielen die ClC Proteine wichtige Rollen bei der Stabilisierung des Membranpotentials, dem transepithelialen Ionentransport und der vesikulären Ansäuerung. Als Alleinstellungsmerkmal der ClC Familie kommen in dieser sowohl chloridspezifische Kanäle als auch sekundäraktive Chlorid/Proton Antiporter vor. Dieser funktionalen Vielseitigkeit steht ein konservierter, gemeinsamer Aufbau gegenüber, welcher aus einer transmembranen Transportdomäne und einer zytoplasmitschen Regulationsdomäne besteht.

Die zytoplasmatischen Domänen eukaryotischer ClC Proteine sind von entscheidender Bedeutung für deren Funktion und sind an der funktionalen Regulation des Chloridtransports beteiligt. Sie wurden mit Vorgängen wie der Regulierung der Kanalaktivität durch Ligandenbindung in Verbindung gebracht. Man geht davon aus, dass diese Regulation durch die Interaktionen zwischen den zytoplasmatischen Domänen des homodimeren Proteins bewirkt wird. Daher war es das Ziel dieser Dissertation, den oligomeren Aufbau der zytoplasmatischen Domänen aufzuklären. Die Struktur der zytoplasmatischen Domäne des humanen ClC-Ka Kanals, welche im Rahmen dieser Arbeit bestimmt wurde, hat sich dabei als wertvoll erwiesen. Die Kristallstruktur der zytoplasmatischen Domäne des humanen ClC-Ka Kanals wies eine ungewöhnliche Art der Dimerisierung auf. Durch die Kombination von Mutagenese, chemischer Querverknüpfung und analytischer Ultrazentrifugation konnte gezeigt werden, dass diese ungewöhnliche Dimerisierung auch in Lösung vorliegt. Die entsprechenden ergänzenden Experimente mit der zytoplasmatischen Domäne von ClC-0 zeigten, dass diese Domäne einen vergleichbaren strukturellen Aufbau besitzt und dass diese Wechselwirkungen sehr wahrscheinlich auch beim vollständigen ClC-0 Protein auftreten. Auf diese Weise konnte der zuvor nur uneindeutig bekannte oligomere Aufbau der zytoplasmatischen Bestandteile der eukaryotischen ClC Proteine bestimmt werden.

Um die Konformationsänderungen der zytoplasmatischen Domäne im Zusammenhang mit der Transmembrandomäne zu verstehen, benötigt man die Struktur eines vollständigen ClC Proteins. Derzeit kennt man jedoch nur die Strukturen von isolierten Transmembran- und zytoplasmatischen Domänen der ClC Familie. Daher war es ein weiteres Ziel dieser Dissertation die Überproduktion und Reinigung eines eine zytoplasmatische Domäne enthaltendes ClC Proteins zu etablieren. Dieses sollte als

Modellsystem für die strukturelle und funktionale Charakterisierung dienen. Zu diesem Zweck wurde eine Reihe prokaryotischer ClC Homologe mit zytoplasmatischer Domäne bestimmt und mit Hinblick auf Expression in *E. coli*, Reinigung und Kristallisation untersucht. Nach genauer Überprüfung wurde RmlClC, ein ClC Homolog aus dem Bakterium *Ralstonia metallidurans*, als einziges für Kristallisationsexperimente geeignete Protein ausgewählt. Trotz breit gefächerter Kristallisationsversuche mit dem vollständigen RmlClC konnten keine Kristalle erhalten werden. Wahrscheinlich liegt dies an einer beweglichen Verbindung zwischen der Transmembran- und der zytoplasmatischen Domäne. Daher wurden diese Domänen getrennt exprimiert und kristallisiert. Mit den nur schlecht beugenden Kristallen der Transmembrandomäne konnte keine detaillierte Struktur bestimmt werden. Die Struktur der zytoplasmatischen Domäne konnte jedoch mit einer Auflösung von 1.6 Å bestimmt werden und zeigt im Detail den strukturellen Aufbau dieser Domäne in prokaryotischen ClC Proteinen. Obwohl diese Domäne eine typische CBS Faltung aufweist, liegt nicht die in Eukaryoten konservierte dimere Anordnung vor. Es konnte gezeigt werden, dass die zytoplasmatische Domäne von RmlClC in Lösung als Monomer auftritt, was möglicherweise auf eine andere Anordung der zytoplasmatischen Domänen in prokaryotischen ClC Proteinen hinweist.

Die in dieser Arbeit vorgestellten Untersuchungen beschreiben die strukturellen und biochemischen Eigenschaften des oligomeren Aufbaus der zytoplasmatischen Bestandteile in eukaryotischen und prokaryotischen ClC Proteinen. Diese Ergebnisse werden als Grundlage für zukünftige Untersuchungen dienen, welche möglicherweise die Mechanismen der Regulierung durch die zytoplasmischen Domänen von ClC Chloridkanälen und –transportern aufzeigen.

# 1 Introduction

## 1.1 Principles of ion transport

All living cells are surrounded by membranes that separate the cytoplasm from the extracellular environment. In eukaryotic cells, membranes in addition compartmentalize the intracellular organelles. Membranes are composed of amphiphilic phospholipids. In these molecules hydrophilic head groups that contain a phosphate moiety and hydrophobic tails of two fatty acid side chains are bound as esters to a central glycerol molecule. Phospholipids spontaneously form lipid bilayers in water, where the head groups are exposed to the aqueous environment and the tails form a hydrophobic core. The dielectric constant of the hydrophobic core is around 2, which is drastically lower than the dielectric constant of water ($\varepsilon_{water}$=80). Due to this property membranes serve as barriers for hydrophilic and particularly charged molecules and they behave like electric capacitors that create and conserve an electric membrane potential.

Specific proteins associated with the membrane have evolved to catalyze the transport of different solutes like metabolites and ions through the lipid bilayer. These membrane transport proteins are also amphiphilic, having a hydrophilic surface that is in contact with the aqueous solution or the polar group of phospholipids, and a hydrophobic belt that interacts with alkyl side chains of lipids in the core of the membrane.

The selective transport of ions across biological membranes is an essential feature of living cells. The movement of potassium, sodium, calcium and chloride underlines important processes like electrical signaling in neurons and muscles, cell volume regulation and transepithelial transport. The proteins that facilitate the transport of ions across cell membranes are divided in two classes, channels and transporters, based on the difference in the thermodynamics of their transport mechanism. Independent of the mechanism, both channels and transporters provide ions with hydrophilic interactions to mimic the dielectric environment of water during the transport through the membrane.

**Ion channels** act as passive conduits through which ions rush down their electrochemical gradient. Ion flow through channels generates transmembrane electric currents that can cause changes in membrane potential. Two fundamental properties of ion channels are ion selectivity and gating. Selectivity refers to the ability of some channels to discriminate between ion species, allowing some to pass through the pore while excluding the others. This discrimination takes place at the narrowest part of the pore known as

selectivity filter. Gating is the process of transition between the open and the closed state of a channel. The gating can be regulated by binding of intracellular or extracellular ligands (ligand-gated channels), changes in membrane potential (voltage-gated channels), and changes in temperature or mechanical stress. A structural entity that acts as a physical barrier to ion movement and renders the channel closed is called a gate.

**Ion transporters** move ions against their concentration gradient by coupling ion translocation to a secondary driving force like the hydrolysis of ATP (primary active transporters) or a downhill movement of another ion species (secondary active transporters). Typical primary active transporters are ion pumps like $Na^+/K^+$ ATPases and $Ca^{2+}$ ATPases. Secondary active transporters can be divided into cotransporters and antiporters, according to the relative direction of the coupled downhill and uphill ion flows. During their catalytic process, ion transporters work by an alternate access mechanism that requires at least two distinct protein conformations. The two conformations expose ion-binding sites to either side of the membrane and often have a high affinity for the ions on one side of the membrane and low affinity on the other side of the membrane. In some cases, there is a defined occluded state during the transport cycle where the ions are completely enclosed by the transporter. Although, ion selectivity and gating are terms mostly related to ion channels, these properties are important for ion transporters too. Regarding gates, whereas channels need only a single gate, transporters need at least two gates that should never be open at once (Figure 1.1).

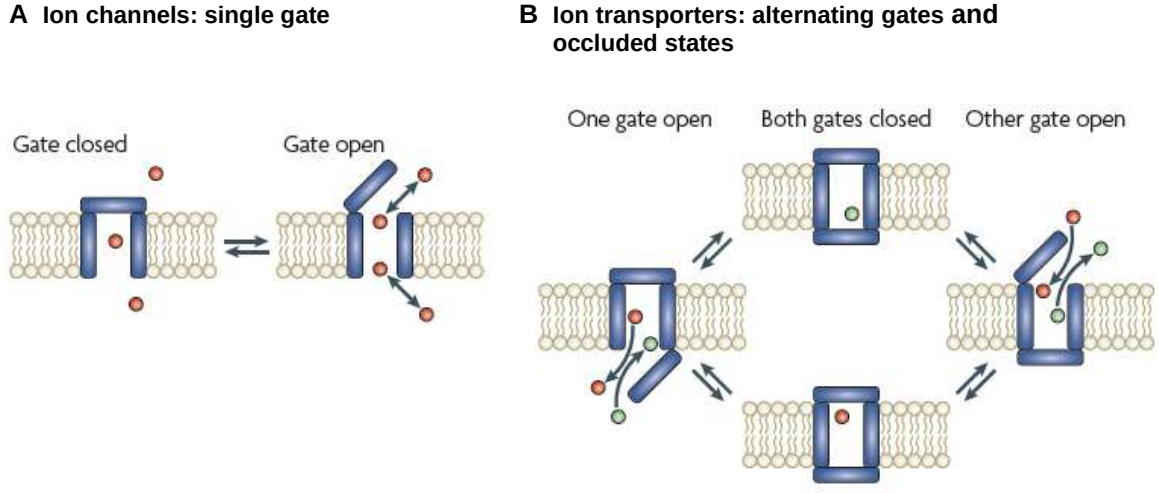



**Figure 1.1. Ion channels versus ion transporters.**
**(A)** Schematic representation of an ion channel as a membrane-spanning pore through which ion (red spheres) movement is controlled by a single gate, represented by a hinged door. **(B)** Ion transporter as a membrane-spanning pore with two gates that open and close alternately. Occluded states, with both gates closed around bound ions, preclude inadvertent opening of the second gate before closure of the first gate, which would allow ions to flow down their electrochemical potential gradient several orders of magnitude faster than the pump can move them against that gradient. Pumped ion movement is rate limited by the gating reactions, rather than by electrodiffusion. Figure obtained from (Gadsby 2009).

Due to their opposed tasks of downhill high-speed ion movement and uphill low-speed ion movement, channels and transporters have traditionally been viewed as completely different entities. Certainly, there is unquestionable evidence for the existence of clear-cut ion channels, like potassium channels, and prototypical ion transporters, like P-type ATPases. However, recent atomic resolution X-ray structures and high-resolution functional measurements have suggested that ion channels and transporters do not need to be as different as initially thought. The remarkable example for blurring this boundary is the ClC family of chloride transport proteins that includes representatives of both channels and transporters.

### 1.1.2 Chloride transport

Chloride is the most abundant anion in nature. Its intracellular concentration is usually low, ranging from 5 mM up to 30 mM, whereas the extracellular concentration is high and reaches 150 mM. Because of its abundance, chloride is the main permeant anion. It is transported by several distinct families of chloride transporting proteins including ClC channels and transporters, ligand-gated GABA and glycine receptors, cystic fibrosis conductance transmembrane regulators (CFTR), Cation Chloride Cotransporters and $Cl^-/HCO_3^-$ exchangers. These proteins reside both in plasma membrane and membranes of intracellular organelles and contribute to a variety of essential processes like the regulation of electrical excitability, transepithelial transport, ion homeostasis and cell volume regulation (Jentsch, Stein et al. 2002). Among those, the ClC proteins encode for the largest family of chloride transport proteins.

## 1.2 ClC chloride channels and transporters

### 1.2.1 History

The discovery of ClC chloride channels dates back to the early 1980s when a chloride channel from the electric organ of *Torpedo californica* was functionally reconstituted and characterized (White and Miller 1979). The name ClC originated from the term $Cl^-$ channel. Ten years later, expression cloning revealed a sequence of ClC-0, the voltage gated chloride channel from *Torpedo marmorata* (Jentsch, Steinmeyer et al. 1990). Shortly after, nine ClC homologs in humans have been identified and characterized by heterologous expression techniques (Steinmeyer, Ortland et al. 1991; Thiemann, Grunder et al. 1992; Kieferle, Fong et al. 1994; Brandt and Jentsch 1995; Steinmeyer, Schwappach et al. 1995). Homology screening and data base searches soon revealed that ClC family

members can be found in all phyla, from bacteria to humans. ClC proteins in model organisms like yeast *Saccharomyces cerevisiae* and nematode *Caenorhabditis elegans* yielded important insights into ClC function (Schwappach, Stobrawa et al. 1998; Schriever, Friedrich et al. 1999; Strange 2003). The breakthrough in the understanding of ClC structure originated from the structure determination of two bacterial homologs (Dutzler, Campbell et al. 2002). The crystal structures allowed for the first time to visualize the complex topology of the membrane embedded catalytic domain and the architecture of the ion translocation pore of this protein family (Dutzler, Campbell et al. 2003). A detailed functional analysis of the crystallized bacterial homologs revealed that they are not chloride channels as previously supposed, but chloride-proton antiporters (Accardi and Miller 2004). This unexpected finding motivated closer examination of the eukaryotic ClC proteins leading to the discovery that some mammalian ClCs also function as $Cl^-/H^+$ antiporters (Picollo and Pusch 2005; Scheel, Zdebik et al. 2005).

### 1.2.2 Overview of the family of ClC proteins

The nine ClC family members in humans can be grouped into three branches by homology. Members of the first branch (ClC-1, -2, -Ka and -Kb) perform their function in the plasma membrane, while the members of the two other branches (comprising ClC-3, -4, -5, and ClC-6 and -7, respectively) are predominantly found in membranes of intracellular vesicles of the endosomal-lysosomal pathway. The proteins that reside in the plasma membrane function as gated chloride channels and play a role in the maintenance of the membrane potential and in epithelial ion transport. The intracellularly located ClCs likely all function as chloride-proton antiporters and are important for the vesicular acidification.

Most of the current knowledge about the physiological functions of mammalian ClCs has been obtained from animal disease models, human genetic diseases and knockout mice (Jentsch, Stein et al. 2002). The findings from these experiments indicate diverse roles for ClC-mediated ion transport.

ClC-1, the closest homolog of ClC-0 from the *Torpedo* electric organ, is expressed almost exclusively in skeletal muscle. Since about 80 % of the ion conductance in resting muscles is provided by ClC-1, the protein is responsible for the stabilization of the resting membrane potential of muscle cells and repolarisation of the membrane after action potentials (Steinmeyer, Ortland et al. 1991). The mutations in ClC-1 lead to mytonia, a form of muscle stiffness due to muscle hyperexcitabilty (Koch, Steinmeyer et al. 1992).

ClC-2 is very broadly expressed (Thiemann, Grunder et al. 1992). This inwardly rectifying $Cl^-$ channel is regulated by a variaty of different stimuli like hyperpolarization, cell swelling or extracellular acidification. It is thus believed to be involved in various

physiological processes. ClC-2 knockout mice, which showed a degeneration of the testis and the retina, suggest a function in the ion homeostasis of these organs (Bosl, Stein et al. 2001).

Two highly homologous channels ClC-Ka and ClC-Kb are both expressed in certain epithelial cells of the kidney and the inner ear, where they have important role in transepithelial ion transport (Kieferle, Fong et al. 1994). They were the first ClC proteins for which a regulatory subunit was discovered (Estevez, Boettger et al. 2001). The β-subunit, Barttin, is needed for the transport of the ClC-K channels to the plasma membrane, but also for the modulation of their physiological function (Scholl, Hebeisen et al. 2006). Mutational inactivation of ClC-Kb channel and Barttin leads to different forms of Bartter syndrome, which is associated with severe renal salt wasting and congenital deafness (Hebert 2003). The disruption of ClC-K1 in mice leads to diabetes insipidus (Matsumura, Uchida et al. 1999).

The roles of intracellular ClCs are best understood for ClC-5 and ClC-7. The transporter ClC-5 is predominantly expressed in the proximal tubule of the kidney, where it localizes to apical endosomes (Gunther, Luchow et al. 1998). ClC-5 co-localizes with the vesicular $H^+$-ATPase and it is presumed that it dissipates the electric potential created by the ATPase during acidification of endosomes. Malfunction of ClC-5 leads to Dent's disease, an X-chromosome-linked disorder associated with proteinuria and kidney stones. ClC-7 is ubiquitously expressed and localizes prominently in lysosomes and late endosomes (Brandt and Jentsch 1995). Additionally, it is also found in the ruffled border of bone-resorbing osteoclasts, a specialized plasma membrane that faces the resorption lacuna. Similar as ClC-5, ClC-7 co-localizes with the V-type $H^+$-ATPase, and is therefore thought to be involved in efficient acidification of vesicles and lacuna. Mutations in ClC-7 are the cause of several variants of osteopetrosis, resulting from impaired bone resorption (Kornak, Kasper et al. 2001). It has recently been discovered that ClC-7 also interacts with a β-subunit, however, functional implications of this protein Ostm-1 are still unclear (Lange, Wartosch et al. 2006).

` There is currently little data about the physiological function of ClC-3, -4 and -6. The knockout mice of ClC-3 showed severe degeneration of the hippocampus and the retina, but no association with human diseases could have been made so far (Stobrawa, Breiderhoff et al. 2001). ClC-4 is broadly expressed in intracellular membranes; it facilitates endosomal acidification and is important for endocytosis (Mohammad-Panah, Harrison et al. 2003). ClC-6 is expressed in neurons of the central and peripheral nervous system and its disruption leads to lysosomal storage diseases and reduced pain sensitivity (Poet, Kornak et al. 2006).

In summary, ClCs play a diverse role in a variety of physiological processes. However, several of their functions are still only poorly understood. The clarification of

these processes requires a thorough understanding of the behavior of ClC proteins on a molecular level.

### 1.2.3 Molecular architecture and ion conduction in ClC proteins

Despite their functional diversity the ClC proteins share a conserved structural organization consisting of a transmembrane catalytic domain and in many cases of a cytoplasmic regulatory domain. Already from the early functional experiments on ClC-0 it has been proposed that ClCs are homodimers with a gated pore contained in each subunit (Middleton, Pheasant et al. 1994). This double-barreled architecture was confirmed by the structure determination of two bacterial ClC homologs from *Escherichia coli* and *Salmonella typhimurium* (Dutzler, Campbell et al. 2002).

The high resolution structure of the bacterial homolog EcClC provides insight into the transmembrane architecture of the family and into the chemistry of chloride binding (Dutzler, Campbell et al. 2003). The tertiary structure of the catalytic domain of ClC proteins is complex and it is comprised of 18 α helices, 17 of which are embedded in the membrane. Only some of these transmembrane helices span the membrane completely and most are severely tilted with respect to the plane of the bilayer. The topology of the membrane domain of a single ClC subunit is shown in Figure 1.2.

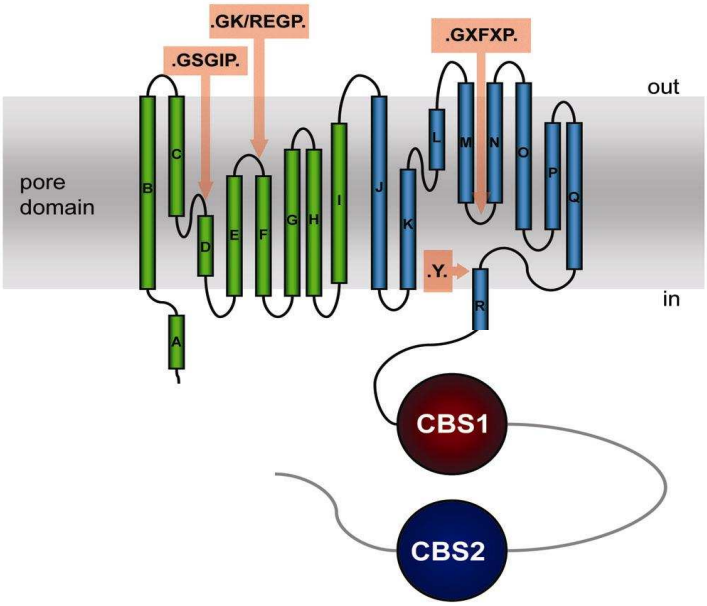



**Figure 1.2. Topology of eukaryotic ClC proteins.**
A schematic view of the ClC subunit containing a transmembrane and a cytoplasmic domain. The topology of the transmembrane domain is inferred from the known structure of EcClC. The 18 α helices are labeled A-R and the two similar halves within the transmembrane domain are colored in green and blue. The sequence regions, which contribute to the $Cl^-$ selectivity filter, are marked with arrows and the respective conserved sequences are shown. The two CBS subdomains of the cytoplasmic part are shown as red and blue spheres.

An interesting feature of the ClC subunit is its antiparallel architecture formed by structurally related N-terminal and C-terminal halves that span the membrane in opposite directions. At the interface between the two halves a translocation path for chloride is found. The structure of EcClC has revealed the residues, which coordinate chloride ions during their passage through the membrane. These residues are located in dispersed regions in the primary sequence of the protein and are highly conserved among ClC channels and transporters. As identified in the structure they create three selective binding sites for chloride, which were termed $S_{ext}$, $S_{cen}$ and $S_{int}$ based on their location along the translocation pathway (Figure 1.3B). In the wild type protein two of the binding sites ($S_{cen}$ and $S_{int}$) are occupied by $Cl^-$ ions, whereas the third site at the extracellular end of the selectivity filter is occupied by the side chain of a conserved glutamate residue (E148). When mutating this residue to glutamine, the binding site is released and binds a $Cl^-$ instead, while leaving the rest of the filter unchanged (Dutzler, Campbell et al. 2003). All three sites can bind chloride with similar affinities in the mM range as determined by crystallography (Lobet and Dutzler 2006). Recent biochemical experiments additionally revealed that the central binding site plays the most important role in chloride selectivity in EcClC (Picollo, Malvezzi et al. 2009).

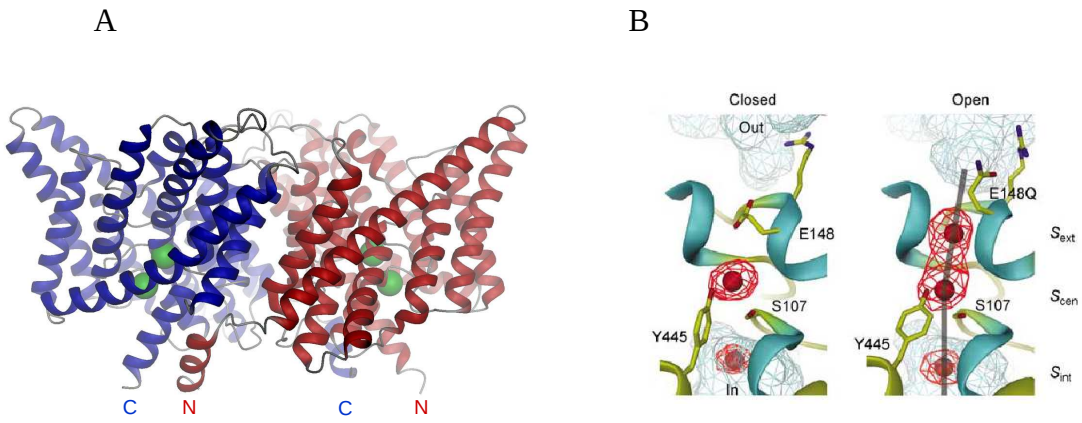



**Figure 1.3. Structure of EcClC and its selectivity filter.**

(A) A side view of the EcClC dimer. Two subunits are colored in red and blue and bound ions in the selectivity filter of each subunit are shown as green spheres.

(B) Selectivity filter of the wild type protein and of selectivity filter mutant. The protein backbone is shown as ribbon with the N-termini of α-helices colored in blue. Selected side chains are shown as sticks, and bound ions as red spheres. Aqueous cavities approaching the selectivity filter from both sides of the membrane are shown as blue mesh. The ion binding sites are labeled. Left: the selectivity filter in a closed conformation - two chloride ions are bound to the sites $S_{int}$ and $S_{cen}$, whereas the ion binding site $S_{ext}$ is occupied by the side chain of the conserved glutamate residue E148. Right: the selectivity filter of the mutant E148Q in an open-like conformation. The mutated side chain of the residue has moved out of the filter and points towards the extracellular solution. An additional ion has taken its place, binding to the site $S_{ext}$. The filter is occupied by three ions that bind in a single file and bridge the aqueous solutions on both sides of the membrane. Figure derived from (Lobet and Dutzler 2006).

It was proposed that structure and ion binding properties of ClC channels and transporters are similar and that even the mechanisms of permeation and gating in channels and coupled transport in transporters have some common features (Miller 2006). Therefore, the molecular architecture of the ion translocation pore observed in EcClC serves as a valuable model for explanation of the mechanistic properties in both ClC channels and transporters.

### 1.2.4 ClC chloride channels

The key properties of ion channels, selective conductance and gating, have for ClC proteins been mainly revealed from electrophysiological experiments on the muscle type channels, ClC-0 and ClC-1. A reason for that is the high conductance that these channels show in comparison to other family members. The single channel conductances of ClC-0 and ClC-1 are 10 pS (Miller and White 1984) and 1 pS (Pusch, Steinmeyer et al. 1994) respectively. The EcClC transporter, in contrast, transports only 2000 ions/s which is about 5 orders of magnitude slower than observed for chloride transport in ClC-0.

Since chloride is the most abundant permeant anion under physiological conditions, selectivity among different anions might not be of large physiological importance. ClC proteins thus are found to be permeant to different small monovalent anions with a preference for $Cl^-$ (i.e. the selectivity for ClC-0 is $Cl^- > Br^- > NO_3^-$ (Rychkov, Pusch et al. 1998)).

Gating of ClC channels has been most extensively studied in ClC-0. Single channel recordings on the protein reconstituted in artificial lipid membranes have revealed some characteristic features (Figure 1.4). First, the equally spaced conductance states that are binominally distributed hinted at a protein with two identical and independent pores (Miller 1982). Second, a complex regulatory behavior includes individual gating acting on each pore separately and common gating affecting both pores simultaneously. Due to their kinetics in ClC-0 they have been also termed fast and slow gating, respectively.

Individual gating is controlled by pH, chloride and voltage. Opening is favored by low pH, a high external chloride concentration and depolarizing voltage (inside positive). The voltage dependence of the individual gating is caused by the movement of a $Cl^-$ ion into the transmembrane electric field during the opening step (Pusch, Ludewig et al. 1995). Therefore, there is a link between the chloride concentration and membrane voltage in opening the channel. Lower extracellular $Cl^-$ concentrations shift the voltage dependence of the open probability to more positive potentials and like this render opening more difficult. This mechanism of gating by permeant ions is profoundly different from classical voltage-gated cation channels, where the gating charge is located on the protein in form of

several positively charged residues in a “voltage sensor” domain (Long, Campbell et al. 2005). The structural determinant of the individual gating process has been identified to be a single glutamate residue that either occupies the extracellular binding site for chloride, thereby prohibiting ion flow in the closed state, or moves away from this site allowing chloride to diffuse through the selectivity filter in the open state of the channel (Figure 1.3B).

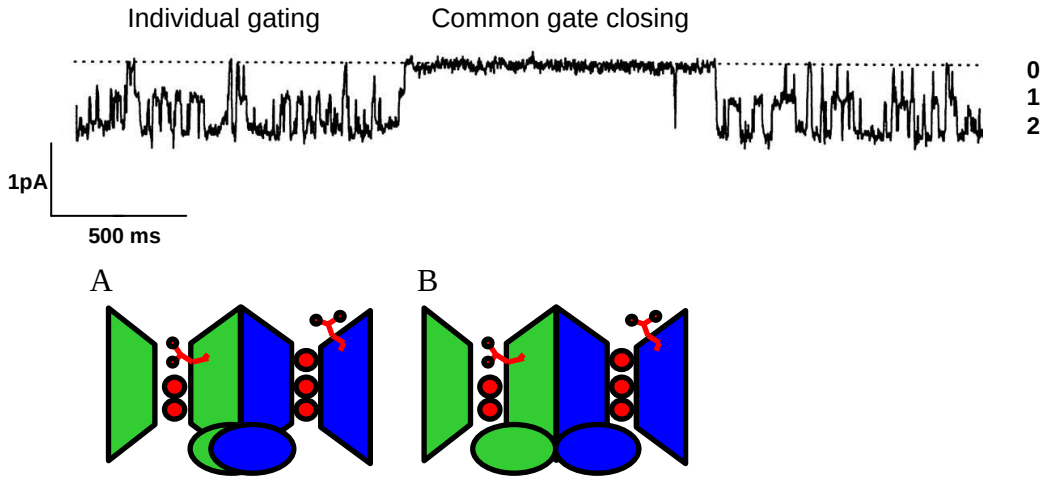



**Figure 1.4. Gating in ClC-0.**
The upper panel shows single channel current trace of ClC-0. Three conductance states are indicated: closed (0), one pore open (1) and two pores open (2). The lower panels A and B schematically represent some of the conformations in the protein during the gating process. The two subunits with independent pores are colored in green and blue and the gating glutamate is shown. Chloride ions are colored in red. Bursts of activity with rapid fluctuations are correlated with the opening and closing of the individual gates. Temporary loss of conduction is associated with the closing of the common gate that blocks both pores simultaneously. The common gating is thought to include the movement of the cytoplasmic domains (green and blue ovals).

The mechanism of common gating, acting on the two individual pores simultaneously, is less well understood. The common gate is influenced by the chloride concentration, temperature, pH and voltage. Membrane depolarization favors common gate closing in ClC-0, while in ClC-1 it favors opening (Fong, Rehfeldt et al. 1998; Accardi and Pusch 2000). Several experiments have suggested conformational changes in the channel underlying the concerted closing of both pores during common gating (Pusch, Ludewig et al. 1997; Bennetts, Roberts et al. 2001, Bykova, Zhang et al. 2006). Additionally, mutations in the dimer interface of the catalytic domains (Duffield, Rychkov et al. 2003) as well as in the cytoplasmic domains have been described to influence the common gate (Fong, Rehfeldt et al. 1998). The indications for an involvement of the intracellular segments of the ClCs in this regulatory process are discussed in the following chapters.

### 1.2.5 ClC transporters

The bacterial homolog EcClC, whose structure has been determined, was the first ClC that has been discovered to function as chloride proton antiporter (Accardi and Miller 2004). Although the detailed mechanism of exchange is unknown, several of its basic features are established.

EcClC transports its substrate ions with a strict stoichiometry of 2 $Cl^-$ for 1 $H^+$. Each subunit acts as an independent transporter with the transmembrane pathways in which chloride and protons move in opposite directions (Nguitragool and Miller 2007). The pathways for $Cl^-$ and $H^+$ overlap on the extracellular side of the protein, but emerge at distant point on the intracellular surface of the protein (Figure 1.5) (Accardi, Walden et al. 2005). The proton pathway cannot be visualized crystallographically at the current resolution of the structures, but functional experiments have identified some of its characteristics. It has been shown that the gating glutamate ($Glu_{ext}$) is the part of the proton path towards the extracellular side and that a corresponding glutamate residue exists on the intracellular side of the membrane ($Glu_{int}$). Mutation of $Glu_{ext}$ in EcClC uncouples $Cl^-/H^+$ transport and converts the protein into a constitutively open chloride conductor (Accardi and Miller 2004). Mutation of $Glu_{int}$ abolishes proton transport but, in contrast to $Glu_{ext}$, it retains the activation by pH. The inner glutamate is conserved within the transporter branch of the ClC family and substituted by hydrophobic residues in the channels.

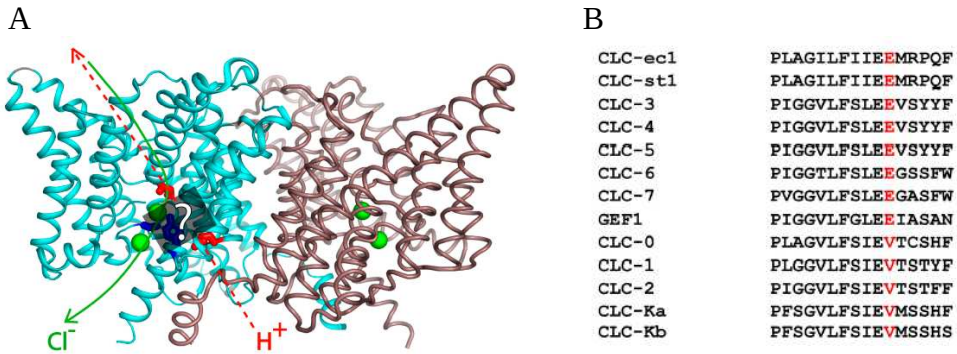



**Figure 1.5. Separate ion pathways in $Cl^-/H^+$ antiporters.**

(A) ClC dimer with the following side chains highlighted: E148, E203 (red); S107, Y445 (blue). $Cl^-$ ions are shown as green spheres. Proposed $Cl^-$ and $H^+$ pathways are indicated in green and red, respectively. The region between E148 and E203, which the $H^+$ must traverse, is indicated by the gray box.

(B) Sequence alignment of the region of $Glu_{int}$ (indicated in red). The residue is conserved in ClC transporters, but not in ClC channels.

Figures derived from (Accardi, Walden et al. 2005).

As previously mentioned, transporters are believed to have two gates, one on each side of the membrane, that alternate their opening and closing in order to prevent free diffusion of the substrates. It was clear from the beginning that the $Glu_{ext}$ plays the role of the extracellular gate in transporters. The intracellular gate, in contrast, was not immediately recognized, but recent experiments have shown that a tyrosine residue (Y445 in EcClC) that is part of the central $Cl^-$ binding site is part of the intracellular gate (Jayaram, Accardi et al. 2008). It has thus been concluded that the structure of WT EcClC represents an occluded state with both gates closed. Removal of both gates by mutations converts EcClC to an uncoupled, ungated channel-like protein with increased conductance.

In the light of the bacterial transporter structure and known mechanistic features of $H^+/Cl^-$ exchange, it has been argued that ClC channels appeared in evolution as "degraded" ClC transporters in which transmembrane proton movement was still coupled to the outer gate at $Glu_{ext}$, but the role of tyrosine residue as an inner gate had been lost (Miller 2006; Lisal and Maduke 2008). This hypothesis would explain the evolutionary relationship between the channels and transporters and the proposed similarities in their mechanisms of action. Nevertheless, a structure of a proper ClC chloride channel is needed to clarify the differences that have to exist in order to support the diverse modes of chloride transport.

## 1.3 Cytoplasmic domains in ClC chloride channels and transporters

Ion channels and transporters frequently have a modular architecture that consists of a membrane bound catalytic domain and a cytoplasmic regulatory domain. To fully understand the mechanism of ion transport and its regulation within the cell, it is critical to define the roles of cytoplasmic domains. The mechanisms by which these domains are coupled to the transmembrane pore, and how they recognize auxiliary subunits and intracellular signals, are some of the questions that are waiting to be answered. The superfamily of potassium channels offers detailed insight into the regulation of channel activity by cytoplasmic domains. The gating of potassium channels can, for example, be regulated by calcium in calcium-sensitive potassium channels or by cyclic nucleotides in cyclic-nucleotide-gated (CNG) and hyperpolarization-activated cation channels (HCN) (Roosild, Le et al. 2004). As part of the regulatory mechanism, cytoplasmic domains can upon ligand binding transiently vary their symmetry, as observed for the cytoplasmic domains of CNG and HCN channels (Zagotta, Olivier et al. 2003).

Cytoplasmic domains are found in all eukaryotic and in a number of prokaryotic ClC proteins. The prokaryotic ClCs of known atomic structure, however, do not have these cytoplasmic components. In eukaryotic proteins the cytoplasmic domains have been shown to be critical for proper protein function (Schmidt-Rose and Jentsch 1997; Maduke,

Williams et al. 1998). Their biological importance is emphasized by a number of mutations in these domains of various family members that lead to diseases like myotonia, Dent's disease, osteopetrosis and classical Bartter syndrome (Jentsch 2008).

Cytoplasmic domains vary in length, in different human ClC family members they range from about 155 to 398 amino acids. The common feature in all the members is the presence of two CBS subdomains (CBS1 and CBS2) that are conserved within the family. The variation is due to the different length of the interdispersed linker connecting the two subdomains and a so called C-peptide following CBS2.

**CBS subdomains.** CBS subdomains are named after cystathionine β-synthase where they were initially described (Bateman 1997), but they are found in several other functionally unrelated families like metabolic enzymes, kinases, channels and transporters (Ignoul and Eggermont 2005). The CBS motifs are typically 60 residues long and have a fold that consists of three beta strands and two alpha helices in a β1 - α1 - β2 - β3 - α2 arrangement. They usually occur as tandem pairs that associate in a pseudo two-fold arrangement via the antiparallel β-sheets of the two subdomains formed by β2 and β3 (Figure 1.6). While the β-strands are tightly interacting, the α-helices are positioned at the periphery of the structure leaving a cleft in between the subdomains (Ignoul and Eggermont 2005).

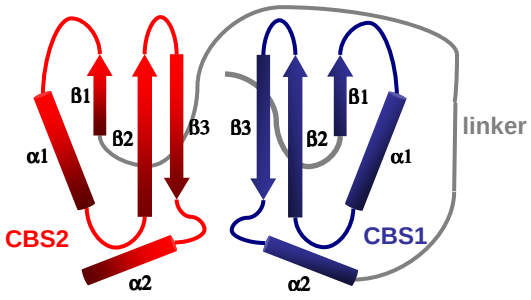



**Figure 1.6. Secondary structure of the CBS domain.**
Two CBS subdomains are colored in blue and red. Each subdomain has a conserved pattern β1-α1-β2-β3-α2. Two subdomains are interacting through the antiparallel β-sheets and leave the α-helices on the periphery.

**Dimerization of CBS domains.** A general feature of CBS domains observed in different prokaryotic and eukaryotic proteins is their ability to dimerize. A number of determined 3D structures of mostly prokaryotic proteins revealed common dimerization modes in which two CBS domains interact via the alpha helices in two possible arrangements: head-to-head or head-to-tail (Figure 1.7 A and B). Assembled dimers have a disc-like appearance with a buried interaction interface of 3000 to 4000 $Å^2$. The self-

association is mediated by the hydrophobic surface of the α1 and α2 helices of both CBS subdomains. Besides the common dimerization modes where both subdomains are in contact, there are examples of CBS domain dimers where the interaction is formed only between α1 and α2 helices of one subdomain, either CBS1 or CBS2. In the structure of TM0892 protein from *Thermatoga maritima* the alpha helices of the CBS2 have a hydrophobic surface and mediate its self-association in a tail-to-tail manner (Figure 1.7C). The alpha helices of CBS1 have, in contrast, a hydrophilic surface and are splayed away from each other.

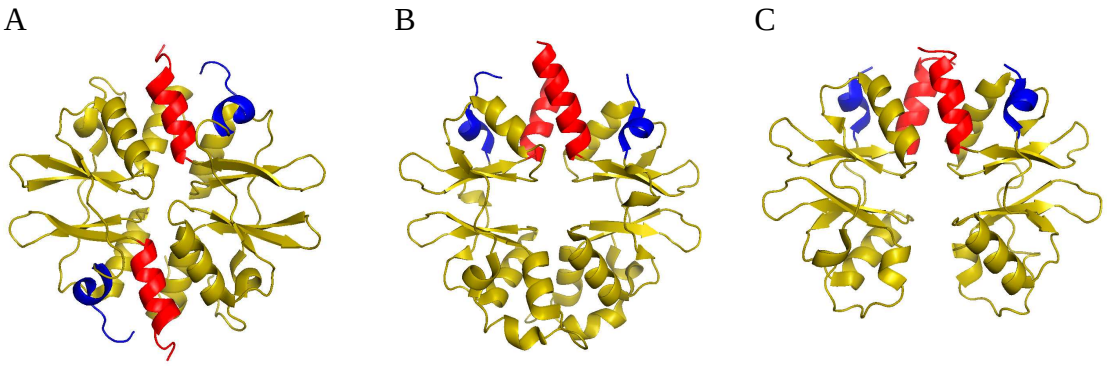



**Figure 1.7. Structures of representative CBS domain dimers.**
The protein chains are shown as ribbon, with the N-terminus colored in blue and the C-terminus colored in red.
**(A)** Structure of the head-to-tail dimer of TM0935 from *Thermatoga maritima* (PDB entry 1O50).
**(B)** Structure of the head-to-head dimer of YkuI from *Bacillus subtilis* (PDB entry 1YAV).
**(C)** Structure of the tail-to-tail dimer of TM0892 from *Thermatoga maritima* (PDB entry 1VR9). There are no interactions between CBS1 subdomains due to its hydrophilic surface.

**Structural insights into CBS domains of ClC proteins.** The structural knowledge on the cytoplasmic CBS domains of ClC proteins at the beginning of this work was obtained from two eukaryotic ClC proteins. The first determined structure was of the cytoplasmic domain of the voltage gated chloride channel ClC-0 from *Torpedo marmorata* (Meyer and Dutzler 2006), which revealed a well-ordered structural core of two tightly interacting CBS subdomains. The 96 residues long linker region between the two CBS subdomains and the C-peptide, in contrast, were not resolved in the X-ray structure. These disordered regions were later investigated by NMR spectroscopy, which has confirmed their large conformational heterogeneity on a fast timescale (Alioth, Meyer et al. 2007). Although the ClC-0 cytoplasmic domain forms dimers in solution, the dimeric arrangement was not found in the crystal. Therefore, a dimerization mode based on the dimeric structure of a prokaryotic CBS domain-containing protein was proposed (Meyer and Dutzler 2006).

The study on the cytoplasmic domain of the human transporter ClC-5 has revealed several novel structural aspects. First of all, the crystal structure of the ClC-5 CBS domain

revealed specific binding of adenosine nucleotides, which was subsequently quantified by binding assays (Meyer, Savaresi et al. 2007). Furthermore, the crystal structure of the ClC-5 domain contained an unusual mode of dimerization for CBS motif containing proteins. In this dimeric arrangement the interface is formed predominantly between the CBS2 subdomains of both chains placing the non-interacting CBS1 subdomains at a larger distance (Figure 1.8). The interaction buries 1600 $Å^2$, which is considerably smaller than in the head-to-head and head-to-tail dimers of other CBS domains.

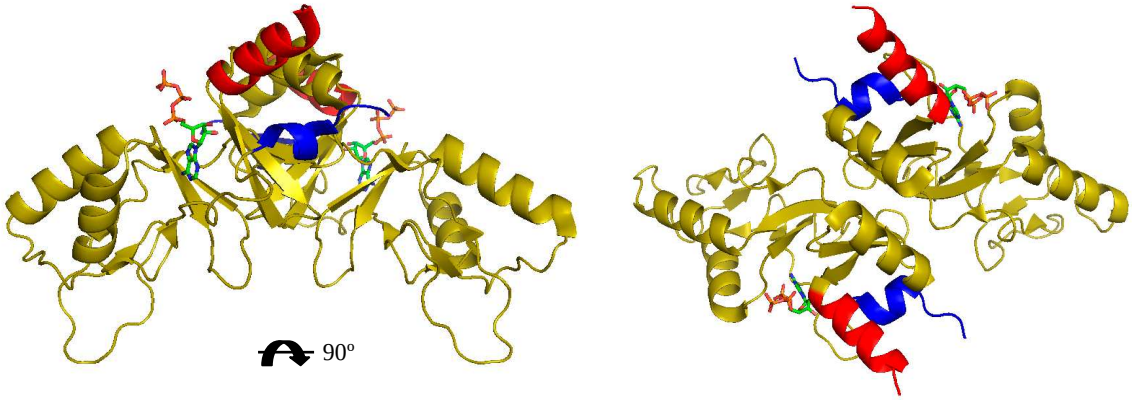



**Figure 1.8. Dimeric arrangement of the ClC-5 CBS domain.**
The protein chains are shown as ribbon, with the N-terminus colored in blue and the C-terminus colored in red. The ATP molecules are presented as sticks. The relationship between the two views is indicated. The V-shaped dimer is formed by the interaction between the CBS2 subdomains.

Both the previously proposed interaction mode for the ClC-0 domain and the one observed in ClC-5 are possible models for the quaternary organization of the cytoplasmic component of ClC proteins. Therefore, the intersubunit organization remained ambiguous and needed to be clarified by further studies.

The cytoplasmic domains in prokaryotic ClC proteins also consist of CBS motifs. The main difference to their eukaryotic counterparts is a smaller size: they are usually not longer than 150 amino acids. This feature is due to the short linker region between two CBS subdomains and a missing or a very short C-peptide. The degree of sequence similarity between eukaryotic and prokaryotic cytoplasmic domains is low, but this is a general characteristic for all CBS domains (Lucas, Encinar et al. 2009). The cytoplasmic domain of the prokaryotic ClC homolog from *Picrophilus torridus* has been crystallized and biochemically investigated (Looser 2008). The structure has a typical CBS domain organization. Both structural and biochemical data showed that this CBS domain, unlike its eukaryotic counterparts, exists as a monomer in solution. The crystallographic analysis of several crystal forms of the PtClC domain did not reveal any interaction in the crystals that

resembles head-to-head and head-to-tail dimers, or dimers found for eukaryotic ClC cytoplasmic domains. The only two-fold arrangement observed in the crystals is presented in Figure 1.9. A similar behavior was later found for the cytoplasmic domain of the ClC homolog from *Ralstonia metallidurans*, which is presented in this work (see Results 2.2.3).

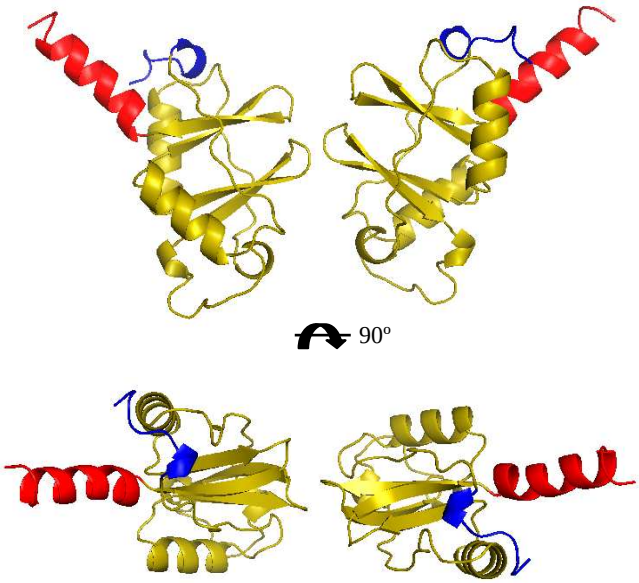



**Figure 1.9. Dimeric arrangement of the cytoplasmic domains of the PtClC homolog found in the $p3_121$ crystal form.** The protein chains are shown as ribbon, with the N-terminus colored in blue and the C-terminus colored in red. The relationship between the two views is indicated. The interaction is formed between the residues of CBS1 subdomains burying the combined surface of 640 $Å^2$.

### 1.3.1 Regulatory role of cytoplasmic CBS domains

Depending on the protein in which they occur, CBS domains have been proposed to have different regulatory roles including the oligomerization and sorting of proteins, channel gating and ligand binding (Ignoul and Eggermont 2005). However, the exact functional purpose for most of the CBS domains in the context of the full length proteins is to this point not fully understood.

A regulatory role has been assigned to the cytoplasmic domains of ClC proteins already in the early studies that included investigations of the effects of deletions and mutations in the CBS domain region on the expression and function of different channels and transporters. Several results suggested an involvement of the intracellular domains in folding and trafficking of ClCs (Schwappach, Stobrawa et al. 1998; Schwake, Friedrich et al. 2001). In addition, an influence of the cytoplasmic domains on the modulation of conduction properties of the ClC channels has been reported (Estevez, Pusch et al. 2004). Several mutations that affect common gating of ClC-1 channel and lead to myotonia have

been identified in its carboxyl tail (Aromataris and Rychkov 2006). These mutations shift the voltage dependence of the common gate to more positive potentials. Although, the common gating of ClC channels is still not well understood, it is likely that it involves significant structural rearrangements of the protein, because the temperature dependence of this type of gating is unusually high (Pusch, Ludewig et al. 1997; Bennetts, Roberts et al. 2001). It also established that common gating involves changes in the interactions between the two subunits and possibly between the cytoplasmic domains. A recent study on ClC-0 confirmed that large movements of the intracellular domains are linked with the common gating of the channel (Bykova, Zhang et al. 2006). By combining "Förster resonance energy transfer" with electrophysiological measurements, it has been discovered that the association of the cytoplasmic domains occurs when the common gate opens, and that separation leads to common gate closure. Although, these experiments revealed novel insight into the structural basis for common gating, they also raised the question how the cytoplasmic events in the C-terminus affect ion permeation in the transmembrane part. The candidate for coupling the C-terminal movement to the pore is the last transmembrane helix R. The position of R helix is unique, since it is partially integrated into the membrane and partially projects into the cytoplasmic space where it connects to the cytoplasmic domain. The tyrosine residue in the N-terminus of R helix is involved in coordinating a chloride ion in the central binding site of the selectivity filter. It appears likely that the movement in the cytoplasmic domains during common gating affects ion permeation of the channel by altering the position of the R-helix. However, to reach a detailed understanding of this process, the structural analysis of cytoplasmic domains in the context of full length ClC proteins is required.

#### 1.3.1.1 Nucleotide binding to the CBS domains

A number of experiments during the last decade revealed that CBS domains can bind adenosine-containing ligands like ATP, AMP and S-adenosylmethionine, which led to the hypothesis that CBS domains function as sensors of intracellular metabolites. The most extensively studied CBS domains are those of the γ subunit of the AMP-activated protein kinase (AMPK) (Scott, Hawley et al. 2004; Rudolph, Amodeo et al. 2007; Townley and Shapiro 2007). The γ subunit contains four repeats of the CBS motifs and can competitively bind AMP or ATP. In this way, the CBS domains of AMPK act as sensors of the cellular energy state. When ATP is depleted due to cellular stress, the kinase is activated by AMP and switches on the catabolic pathways that generate ATP. Different effects of AMP and ATP probably arise from the charge difference between the mono- and triphosphate groups of AMP and ATP that are recognized in the critical regulatory region and transmitted to the catalytic α subunit of AMPK by a still unknown mechanism.

The CBS domains of ClC proteins have also been suggested to constitute binding modules for nucleotides that regulate function. ATP binding has been reported for ClC-1 (Bennetts, Rychkov et al. 2005; Bennetts, Parker et al. 2007; Tseng, Bennetts et al. 2007), ClC-2 (Scott, Hawley et al. 2004) and ClC-5 (Meyer, Savaresi et al. 2007). However, it is still unclear whether nucleotide binding is a universal property of all ClC CBS domains and, if so, to which extent it modulates the function of ClC proteins. For example, it appears that CBS domains of the channels ClC-0 and ClC-Ka lack nucleotide dependence as neither binding assays nor co-crystallization experiments have shown specific interaction of the protein with nucleotides (Meyer and Dutzler 2006; Markovic and Dutzler 2007).

The first direct evidence for nucleotide binding to ClC proteins comes from the structural and biochemical analysis of the ClC-5 cytoplasmic domain (Meyer, Savaresi et al. 2007). The co-crystallization experiments revealed specific binding of ATP, ADP and AMP in the cleft between the two CBS subdomains of ClC-5 (Figure 1.10). From an inspection of the binding pocket the amino acids important for the binding have been detected and a novel conformation of the bound nucleotide was observed. Similar conformation of bound ATP was later found in the equivalent domains of AMPK (Townley and Shapiro 2007). The determined affinities of the nucleotides ATP, ADP and AMP binding to the domain of ClC-5 are very similar in the range of 100 µM.

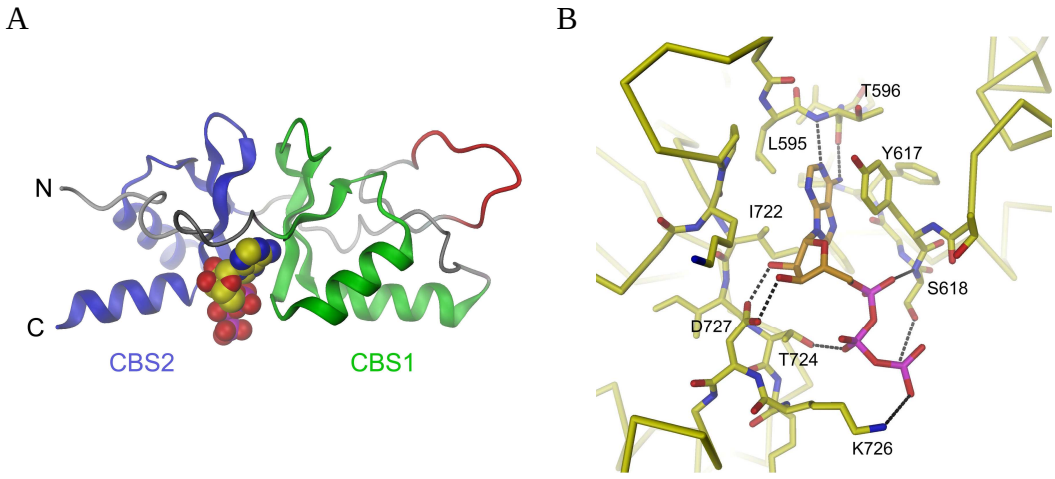



**Figure 1.10. Binding of ATP to the ClC-5 cytoplasmic domain.**
**(A)** Ribbon representation of the ClC-5 domain. The two CBS subdomains are colored in green and blue, respectively, residues of the ubiquitin ligase recognition site are colored in red. The bound ATP molecule is shown as CPK model. **(B)** Structure of the nucleotide binding site. Protein main chain is shown as Cα trace and colored in yellow with selected residues near the bound ATP molecule shown as sticks. Dashed lines denote hydrogen bonds between protein and nucleotide. The protein residues in contact with protein are labeled. The tyrosine residue 617 makes stacking interaction with the base, the aspartate 727 coordinates the ribose and the serine 618 the phosphate group. Mutation of Y617 and D727 abolishes nucleotide binding. Figures derived from (Meyer, Savaresi et al. 2007)

Although the nucleotide binding is abolished by mutating the amino acids that coordinate the ligand, these mutations did not alter the current phenotype of the ClC-5 transporter when studied in heterologous expression systems. Surprisingly, changes in current rectification were observed when such mutations were studied in the background of the uncoupled ClC-5 gating glutamate mutant, which might hint at a regulatory role of nucleotides under currently unknown circumstances. Recently, the intracellular regulation of ClC-5 by adenine nucleotides was directly studied by inside-out patch experiments (Zifarelli and Pusch 2009). The obtained results suggest that nucleotides increase the probability of ClC-5 to be in an active, transporting state.

In contrast to the observations in ClC-5, for the muscle channel ClC-1 it has been proposed that ATP has an inhibitory effect (Bennetts, Rychkov et al. 2005). Detailed electrophysiological analyses have demonstrated that physiological concentrations of ATP in combination with low intracellular pH might inhibit ClC-1 channel by shifting the open probability curve of the common gate towards more depolarizing potentials (Tseng, Bennetts et al. 2007). This would mean that in the presence of ATP and low pH the channel is in the closed state. It is proposed that this mechanism plays an important role in opposing muscle fatigue. Intensive exercise is associated with muscle acidification which has been observed to lead to reduced chloride conduction (Pedersen, de Paoli et al. 2005). Due to reduced $Cl^-$ shunting current on the surface membrane of skeletal muscles partially inactivated $Na^+$ channels are still able to initiate action potential. Therefore, the regulation of ClC-1 is important for the maintenance of excitability in working muscle.

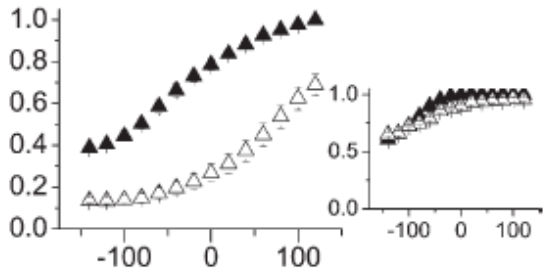



**Figure 1.11. Effect of 1 mM ATP on ClC-1 at pH 6.2.** Po-V curves of the common (left) and fast gate (right). Solid and open lines were obtained in 0 and 1 mM ATP, respectively. ATP shifts the open probability of the common gate by 130 mV towards more positive potentials, but does not influence the fast gate. Figure derived from (Tseng, Bennetts et al. 2007).

The regulation of the ClC-1 common gating is thought to result from a direct interaction between ATP and the cytoplasmic domains, although a structural evidence for the binding is still missing. The residues responsible for the nucleotide binding in ClC-5 are not conserved in ClC-1, but some similarities exist (Figure 1.12). However, the binding hypothesis is supported by the two single-point mutations in the ClC-1 cytoplasmic

domain, H847A and L848A, known to suppress the inhibitory ATP effect (Zhang, Tseng et al. 2008).

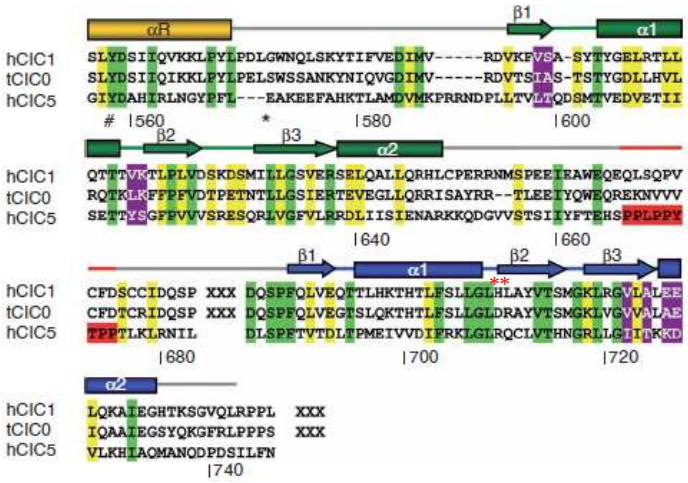



**Figure 1.12. Sequence alignment of the cytoplasmic domains of ClC-1, ClC-0 and ClC-5.** Residues in ClC-5 involved in ATP binding are highlighted in violet. The residues H847 and H848 of ClC-1 are labeled with a red star. Figure derived from (Meyer, Savaresi et al. 2007).

After the structural and functional analyses of ligand binding to the CBS domains of several proteins, the effects these ligands exert on their targets and the molecular mechanisms underlying the regulation are still ambiguous. Currently, there are two different types of regulatory mechanisms proposed (Lucas, Encinar et al. 2009). In the first type, denoted as “static”, the nucleoside position of the ligand induces essentially no change in the protein structure. Instead, the electrostatic potential at the binding site is thought to be the most significant property of adenosine nucleotide binding. This “static” type of regulation would be involved in processes in which regulation by the cellular energy state would be advantageous like, for instance, in AMPK. The second type of mechanism, denoted as “dynamic”, would involve dramatic conformational changes in the protein upon ligand binding. This dynamic type of regulation has recently been observed for the cytosolic domain of the magnesium transporter MgtE (Hattori, Iwase et al. 2009) and for the CBS domain of protein MJ0100 from *Methanocaldococcus jannaschii* (Lucas, Encinar et al. 2009). The cytosolic domains of MgtE have been proposed to regulate the gating of the ion-conduction pore by sensing the intracellular $Mg^{2+}$ concentration. Crystal structures of the cytoplasmic domains in the presence and absence of $Mg^{2+}$ demonstrate a substantial conformational change in the CBS domain dimer upon ion binding (Figure 1.13). This conformational change is transmitted to the ion conduction pore via the plug helix that connects the transmembrane with the cytosolic component. The structural study on the CBS domain of MJ0100 in complex with S-adenosylmethionine and S-methyl-5’-thioadenosine revealed different conformations of the head-to-head dimer depending on the number of ligands bound to the protein (Lucas, Encinar et al. 2009). Four detected conformers differed in the angle adopted between the two CBS domains suggesting an open-to-closed conformational change.

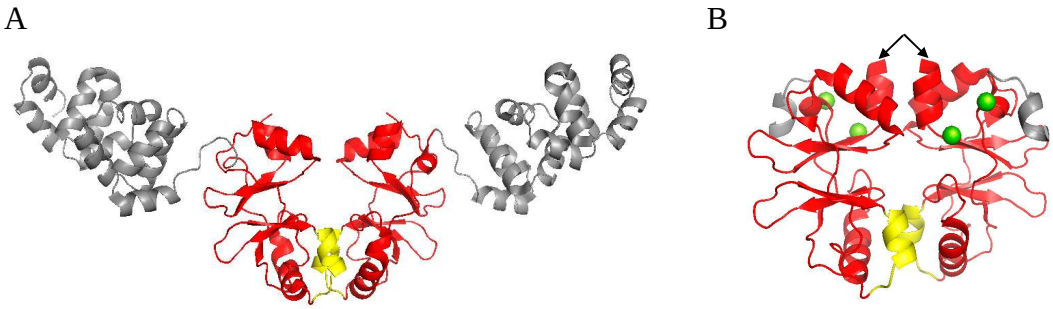



**Figure 1.13. Cytoplasmic domains of MgtE transporter act as ion sensors.**
Structure of the cytoplasmic domains in the absence **(A)** and presence **(B)** of magnesium ions (PDB entries: 2YVZ and 2YVY, respectively). The protein chains are shown as ribbon, with the N-domains colored in gray and CBS domains colored in red. The α2 helices of CBS1 building the interface are colored in yellow, the $Mg^{2+}$ ions are presented as green spheres. In the $Mg^{2+}$ bound structure the N-domains are removed for clarity. Binding of ions influences closing of the CBS dimer. The closing of the ion conduction pore is triggered by the movement of the plug helices (indicated by the black arrows).

It is currently unclear which type of mechanism would be relevant for the ClC proteins ClC-5 and ClC-1 that were identified to interact with nucleotides. The static type of regulation is unlikely due to the fact that both ClC-5 and ClC-1 do not discriminate between ATP, ADP and AMP. Moreover, the binding affinities determined for ClC-5 and the inhibitory effect of the physiological concentrations of ATP on ClC-1 would suggest that the ligands permanently occupy the protein. It could thus be argued that the nucleotides have rather a structural than regulatory role and that most probably CBS domains in ClC proteins do not act as sensors of the cellular energy state. The dynamic type of regulation is more likely, but structural information on CBS domains or full-length ClC proteins in the presence and absence of ligands is necessary to confirm this speculation. The results that support the dynamic type of regulation include the observed structural rearrangements associated with the common gating of ClC-0 and that were also suggested for ClC-1. As nucleotide binding is only shown for certain ClCs it is also possible that other factors (like for example pH in the case of ClC-1) can act as regulators of transport.

## 1.4 Prokaryotic ClC homologs as model proteins

Prokaryotic ClC homologs have served as invaluable model system to shed light on both the structure and function of ClC family members. In addition to providing a high-resolution molecular view of the family, the functional investigation of prokaryotic ClC proteins has led to the discovery that many ClCs are chloride-proton antiporters and not chloride channels. It is thus attractive to assume that additional features of the ClC family that concern concerted regulatory mechanisms could be uncovered by examining

prokaryotic proteins. This goal is facilitated by the recent availability of sequenced prokaryotic genomes that aid the discovery of proteins which have the desired features. A feature that makes prokaryotic proteins attractive targets is their easier overexpression and purification when compared to their eukaryotic counterparts. The prokaryotic expression systems, particularly *E. coli*, are inexpensive, robust, easy to handle and thus suited for broad screening approaches. The availability of milligram quantities of the prokaryotic ClCs in an active form makes them readily accessible for both structural and functional studies.

The newly discovered homologs may prove to be more suitable as model proteins to study certain aspects of ClC function than the structurally characterized EcClC and StClC. Several of them share additional structural features with eukaryotic ClCs and some proteins could potentially function as channels. Sequence analysis of known prokaryotic ClC homologs shows that less than a half of these prokaryotic proteins contain the "intracellular glutamate" ($Glu_{int}$) found in transporters, while the other half contains a hydrophobic residue at this position (Matulef and Maduke 2007). A number of prokaryotic ClCs contain cytoplasmic CBS domains and they are, therefore, very interesting targets for studies on concerted regulatory mechanisms. Many remaining questions are to be addressed on the new prokaryotic ClC homologs: What is the structure and function of the cytoplasmic domains in the prokaryotic ClC homologs? Are there endogenous ligands that bind to these domains? Do any of the prokaryotic ClCs that lack $Glu_{int}$ function as channels rather than transporters? How are the prokaryotic ClCs regulated in their physiological context? Do the prokaryotic homologs have any accessory subunits, such as the two known for mammalian ClCs? Addressing these questions on prokaryotic homologs may reveal novel features that are relevant for the entire family of chloride channels and transporters. However, one has also to be aware that certain regulatory and pharmacological properties will be unique to eukaryotic proteins and that the understanding of such processes will require structural investigation of eukaryotic proteins.

## 1.5 Aim of this thesis

While currently all structural information on the ClC family only exists for the isolated transmembrane catalytic domains and the cytoplasmic regulatory domains, there are no structures of full length proteins that include both components available. Therefore, the main goal of this thesis was to establish the overexpression and purification of prokaryotic ClC proteins containing cytoplasmic domain that would serve as model systems for structural and functional studies. An additional aim of the thesis was to reveal the oligomeric assembly of cytoplasmic domains in prokaryotic and eukaryotic ClC proteins which is of fundamental importance for investigating regulatory mechanisms governed by these cytoplasmic components.

# 2 Results

## 2.1 Expression and biochemical characterization of prokaryotic ClC homologs

### 2.1.1 Homology search

The members of the ClC family of chloride channels and transporters are widely represented in prokaryotes. Over a thousand of sequenced prokaryotic genomes offer, therefore, an enormous source of ClC homologs that could be potentially used for structural studies.

The sequence identity between the prokaryotic and eukaryotic ClCs is relatively low (15 – 25 %). However, several stretches of the ClC sequence are invariant and serve as a good criterion for classification. The highly conserved regions include residues of functional importance that form the ion translocation pore with its specific binding sites.

One interesting structural feature of prokaryotic ClC family members is that some of them contain cytoplasmic CBS domains and some do not. Regarding functional differences, it is known that the ClC homolog from *E. coli*, EcClC, functions as chloride/proton antiporter, while ClC homologs that function as chloride channels have so far not been identified and functionally characterized.

The prokaryotic ClC homologs I have chosen for cloning had the following characteristics:

1) They contain cytoplasmic CBS domain.
2) They either contain a valine, leucine or isoleucine instead of the glutamate at the position determined to be responsible for the proton translocation in ClC transporters on the intracellular side (Accardi, Walden et al. 2005). In that way only potential prokaryotic ClC channels were selected.
3) They are relatively small in size as compact proteins are more promising for crystallization
4) They are found in easily available, non-pathogenic microorganisms

The protein sequence of the chloride channel ClC-0 from *Torpedo marmorata* was used as a query for blast searches (blastp). The Blossum62 matrix was applied for scoring of the alignments. Table 2.1.1 presents 27 homologs that were identified during the search and subsequently cloned. By pairwise sequence alignment implemented in ClustalW, the

| Strain with the strain code | Homolog name | Gene | Identity to ClC-0 [%] | Length [aa] | Mw [kD] |
|---|---|---|---|---|---|
| *Archaeoglobus fulgidus* DSM 4304 | AfClC | AF_1415 | 19 | 589 | 63.4 |
| *Caldivirga maquilingensis IC-167* DSM 13494 | CmClC | Cmaq_0176 | 18 | 586 | 63.2 |
| *Chlorobium limicola* DSM 245 | ClClC | Clim_1107 | 19 | 470 | 50.2 |
| *Chlorobium tepidum TLS* DSM 12025 | CtClC | CT1391 | 16 | 629 | 68.7 |
| *Desulfuromonas acetoxidans* DSM 684 | DaClC | Dace_0926 | 17 | 606 | 65 |
| *Halobacterium sp. NRC-1* | HaClC | VNG1544G | 15 | 792 | 81.7 |
| *Haloferax volcanii* DSM 3757 | HvClC | HVO_A0171 | 15 | 656 | 67.6 |
| *Hydrogenobaculum sp. Y04AAS1* | HyClC | HY04AAS1_0583 | 19 | 462 | 49.6 |
| *Hyperthermus butylicus* DSM 5456 | HbClC | Hbut_0425 | 16 | 563 | 60.1 |
| *Metallosphaera sedula* DSM 5348 | MsClC | Msed_0357 | 15 | 588 | 63.6 |
| *Methanococcoides burtonii* DSM 6242 | MbClC | Mbur_1352 | 14 | 577 | 62.2 |
| *Methanoculleus marisnigri JR1* DSM 1498 | MmClC | Memar_1622 | 13 | 754 | 79.7 |
| *Methanosaeta thermophila PT* DSM 6194 | MtClC | Mthe_1465 | 16 | 580 | 62.2 |
| *Methanosarcina acetivorans C2A* DSM 2834 | MaClC | MA3609 | 17 | 607 | 65.9 |
| *Methanosarcina barkeri* DSM 804 | Mb1ClC | Mbar_A2221 | 16 | 590 | 63.7 |
| *Methanosarcina barkeri* DSM 804 | Mb2ClC | Mbar_A1823 | 16 | 593 | 64.4 |
| *Methanosarcina mazei Go1* DSM 7222 | MmzClC | MM_0498 | 19 | 589 | 63.4 |
| *Pelobacter carbinolicus* DSM 2380 | PcClC | Pcar_1620 | 14 | 606 | 64.7 |
| *Pelodictyon luteolum* DSM 273 | PlClC | Plut_0919 | 15 | 624 | 67.8 |
| *Picrophilus torridus* DSM 9790 | PtClC | PTO1156 | 14 | 590 | 65.4 |
| *Prosthecochloris vibrioformis* DSM 265 | PvClC | Cvib_0969 | 15 | 624 | 68 |
| *Ralstonia metallidurans CH34* DSM 2839 | Rm1ClC | Rmet_0455 | 13 | 560 | 57.8 |
| *Ralstonia metallidurans CH34* DSM 2839 | Rm2ClC | Rmet_0584 | 14 | 637 | 66.8 |
| *Rhodopseudomonas palustris CGA009* | RpClC | RPA1465 | 14 | 604 | 64 |
| *Streptomyces coelicolor A3(2)* DSM 40783 | ScClC | SCO6320 | 16 | 589 | 60 |
| *Syntrophobacter fumaroxidans MPOB* DSM 10017 | SfClC | Sfum_2352 | 15 | 612 | 66.3 |
| *Thermoplasma volcanium GSS1* DSM 4299 | TvClC | TVN0094 | 15 | 592 | 64.2 |

**Table 2.1.1. Prokaryotic ClC homologs.**

**Table 2.1.1. Prokaryotic ClC homologs.**
List of 27 ClC homologs selected for cloning. The homologs were identified using blastp with Blosum62 matrix searching with ClC-0. Identity to ClC-0 was determined applying pairwais sequence alignment implemented in ClustalW.

identity to ClC-0 was determined. The percentage of identical residues between ClC-0 and the prokaryotic homologs is relatively low, in the range between 13 to 19 %.

The size of the selected ClC proteins with cytoplasmic domains varies from 560 to 792 residues (on average 591 residues). The shortest human ClC proteins are ClC-K channels with 687 amino acid residues, which indicates that most of the selected prokaryotic ClCs are more compact proteins due to the shorter loops connecting transmembrane helices.

In searching for putative prokaryotic ClC channels, two proteins that do not contain a cytoplasmic domain were identified: ClClC and HyClC. This was an interesting observation, since all other prokaryotic ClCs lacking cytoplasmic domains are putative transporters, based on the conservation of the intracellular glutamate. Therefore, the two homologs were included in the screen. Figure 2.1.1 summarizes the conservation of the residues constituting the selectivity filter of EcClC in the chosen set of ClC homologs. The structure of the selectivity filter is presented in Figure 2.1.1A. The gating glutamate (E148 in EcClC) is strictly conserved in all ClCs and the surrounding stretch of sequence is highly homologous (Figure 2.1.1C). This is not the case for the residues of the central ion binding site, S107 and Y445 in EcClC. While the tyrosine residue is conserved in most of the selected proteins, the serine is not found in any of them. In more than half of the proteins, the serine residue is substituted by a histidine, and in the rest of the proteins this region is not homologous. As described before, all selected homologs have a hydrophobic residue (mostly valine) at the position of the intracellular glutamate that acts as a proton donor in ClC transporters. Structural information in combination with functional analysis will be the only way to uncover the role of specific residues in the ion binding in this novel set of ClC proteins.

From the multiple sequence alignment a phylogram connecting the 27 ClC homologs with EcClC and ClC-0 was generated (Fig. 2.1.1.B). In phylograms, the distance of one group from the other groups indicates the degree of relationship; i.e. closely related groups are located on branches close to one another. Selected ClC homologs cluster into five different groups. By analyzing the microbial origin of each ClC homolog, it is apparent that the biggest group (blue in Figure 2.1.1B) encompasses proteins coming from thermophilic archaea and δ/ε proteobacteria. Four ClC proteins are from green sulfur bacteria (Chlorobi - green) and five from methanogenic archaea (orange). Two homologs are from halophilic archaea and one from an actinomycete (purple). Another group contains proteins from β- and α- proteobacteria (red). This classification is potentially

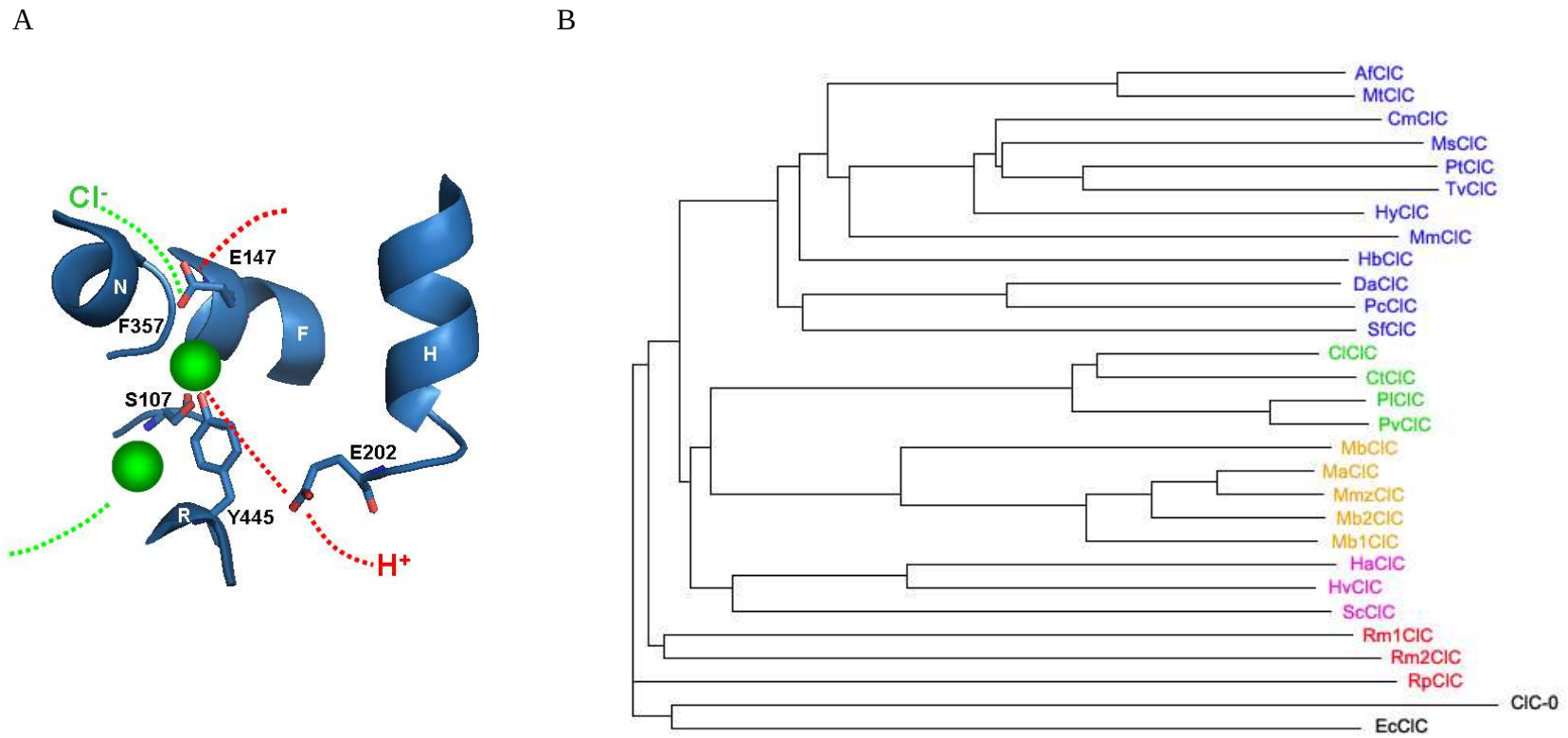
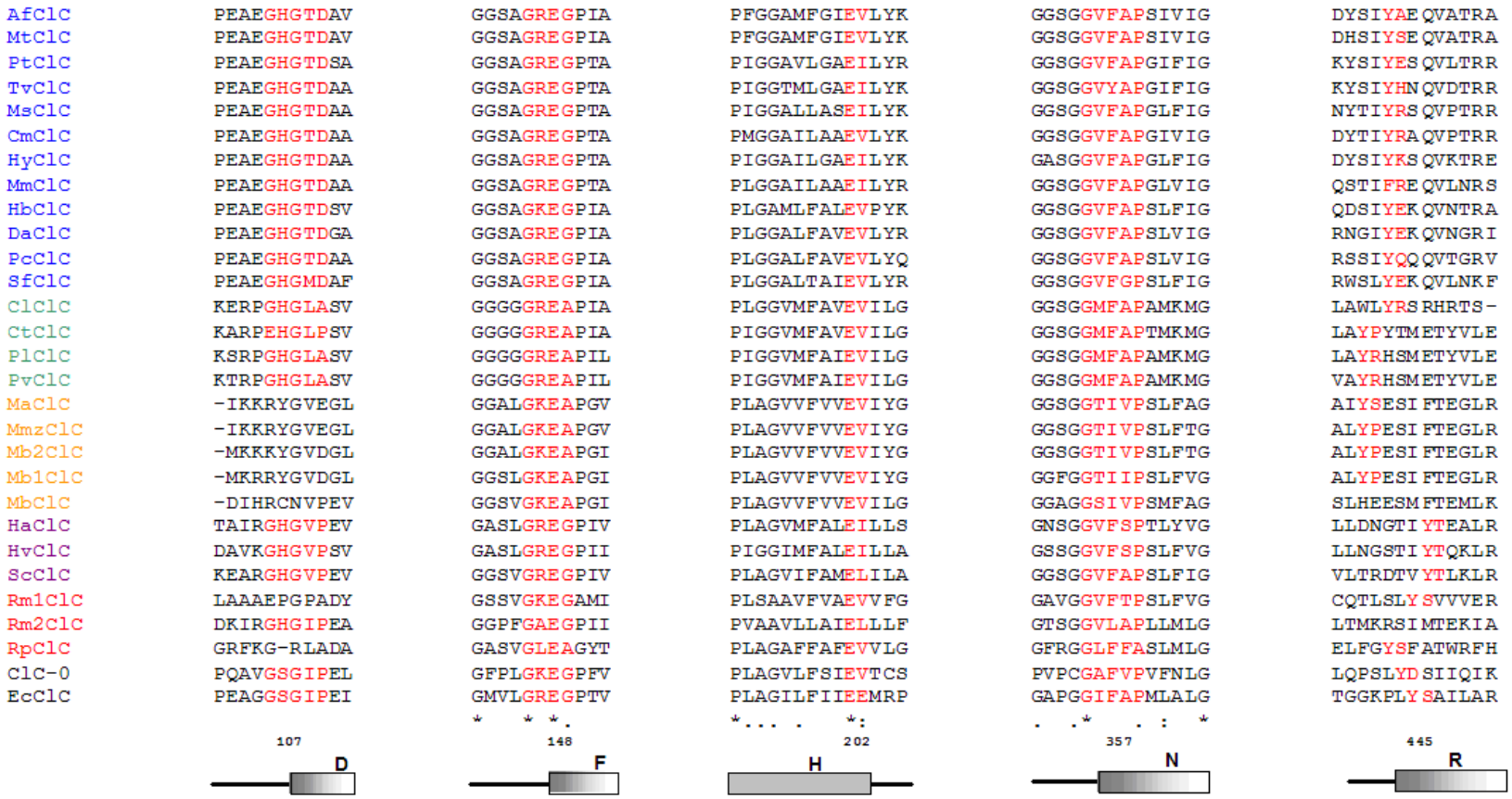



**Figure 2.1.1. Conservation of the residues building the EcClC selectivity filter in the selected prokaryotic ClC homologs.**

A) Structure of the EcClC selectivity filter in the closed conformation. The protein backbone is shown as ribbon, with selected residues as sticks. The chloride ions are represented as green spheres. The residues building the chloride binding sites are labeled, together with the helices they belong to. The putative proton translocation pathway is labeled with the red dotted line connecting the intracellular glutamate E202 ($H^+$ donor) and the extracellular glutamate E148 ($H^+$ acceptor).

B) Phylogram of the ClC homologs created with the multiple sequence alignment in ClustalW. Selected ClC homologs cluster into five different groups based on their sequence similarity i.e. prokaryotic organism they come from. Blue: thermophilic archaea and δ/ε proteobacteria; Green: green sulfur bacteria; Orange: methanogenic archaea; Purple: halophilic archaea and actinomycetes; Red: β and α proteobacteria; ClC-O from *Torpedo marmorata*; EcClC: γ proteobacteria.

C




C) Sequence alignment between selected prokaryotic ClCs, ClC-0 and EcClC (ClustalW). Five sequence regions containing the residues of the selectivity filter are presented. Sequences with high homology are colored in red, identical residues are labeled with a star. The numbering of the critical residues corresponds to EcClC. The gating glutamate is strictly conserved. A serine residue in the selectivity filter (S107 in EcClC) is not conserved and the intracellular glutamate (E202 in EcClC) is substituted with hydrophobic valine, isoleucine or leucine residues.

interesting for correlating the expression and biochemical behavior with the microbial origin of the selected proteins.

### 2.1.2 Cloning of prokaryotic ClC homologs

Selected microbes were purchased from the DSMZ microbial collection as dried cell pellets or as growing cultures. The genomic DNA was subsequently isolated by standard procedure resulting in sufficient starting material for amplification. The genes of interest were amplified by the polymerase chain reaction (PCR), using primers that contained specific sequences for subsequent cloning into expression vectors. The choice of the DNA polymerase had the biggest influence on successful amplification (Pfusion®, KOD® and a mixture of polymerases present in the Master Mix® were used).

For cloning of prokaryotic ClC homologs a new strategy termed fragment exchange (FX) cloning was applied. This type of cloning circumvents the problem of internal cleavage sites for certain restriction enzymes and facilitates high throughput efforts. As for the Gateway® cloning system, the genes of interest and the target vector are mixed in one tube to which the restriction enzyme is added. The enzyme catalyzing the reaction is very specific resulting in a low frequency of cleavage site in the genes of interest that could interfere with the procedure.

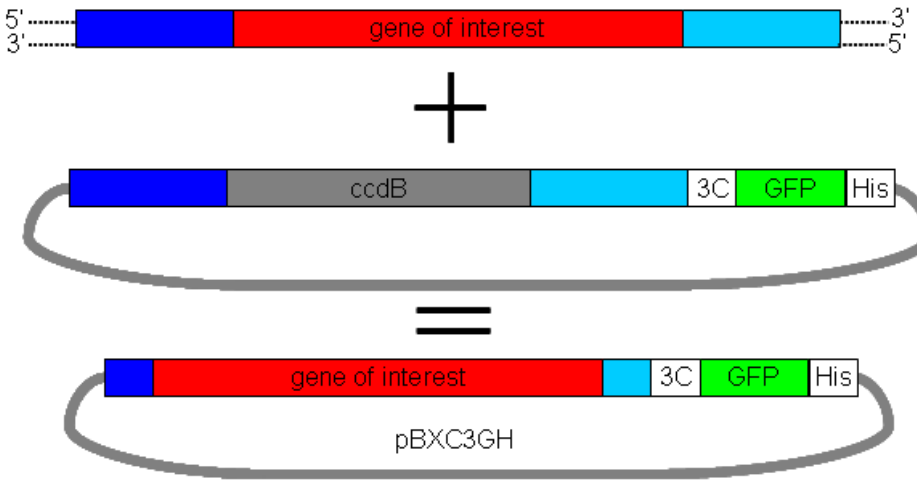



**Figure 2.1.2. FX Cloning**
Fragment exchange (FX) cloning requires the addition of short DNA sequences on both sides of the gene of interest (blue). They differ in sequence to guarantee directional cloning. The target vector contains the CcdB marker (grey) flanked by the corresponding short DNA recognition sequences. After the ligation reaction 3 bp are left on the plasmid introducing an additional amino acid on both sides (5' serine, 3' alanine). CcdB (grey) is a toxic protein to *E. coli* that prevents growth of clones that do not contain the insert.

Another beneficial feature for high throughput cloning is the integration of the selection marker CcdB in the original plasmid (Figure 2.1.2). CcdB is a cytotoxic protein that inactivates DNA gyrase and that is integrated into the plasmid under the control of a consecutive promoter (Bernard, Gabant et al. 1994). The cassette is eliminated after a successful ligation with the insert. In that way, religations with the CcdB cassette and uncut vectors are eliminated since bacteria containing such vectors fail to grow.

All 27 homologs were cloned into pBXC3GH vector to generate C-terminal fusions with the green fluorescent protein (GFP) and $His_{10}$ tag at the end. The GFP part in these fusion proteins can be used as a folding indicator, for detection and quantification of the proteins. Only promising candidates were subsequently cloned into pBXC3H vector which results in a C-terminal fusion with a $His_{10}$ tag used for affinity purification (Figure 2.1.3).

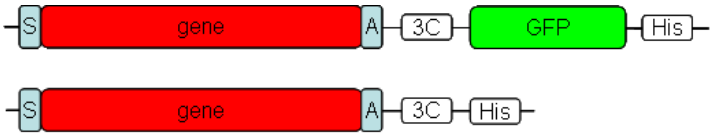



**Figure 2.1.3. Expression constructs for prokaryotic ClC homologs**
Initially, fusion proteins containing C-terminal GFP (green) and His-tag separated from the protein (red) by 3C protease cleavage site were generated by FX cloning (top). Expressing proteins were subsequently cloned into a vector containing a C-terminal 3C site and a His-tag only to aid purification (bottom). The FX cloning method introduces an additional amino acid on both termini (N- serine, C- alanine).

### 2.1.3 Expression tests

The first hurdle towards membrane protein structure determination is the overexpression of target proteins at high levels needed for crystallization. The expressed membrane proteins must have suitable biochemical properties like proper folding, catalytic activity and broad detergent stability to be suitable for structural analysis. When screening broadly for microbial homologs of a certain membrane protein family, the chance for identifying candidates that satisfy the above mentioned criteria is increased. Nevertheless, it is advantageous to be able to initially investigate a large number of constructs in parallel in small volumes to identify the most promising ones for scale-up closer characterization.

Several parameters can influence the amount of well folded membrane protein such as the expression strain, induction temperature, growth medium and promoter strength (Wang, Safferling et al. 2003). Analysis and optimization of expression conditions for a large number of homologs can be laborious when applying standard methods like Western blot or activity assays. Moreover, due to the hydrophobic nature of membrane proteins, their transfer to the blotting membrane can be inefficient, thus making Western blotting

less suitable for quantitative purposes. Since simple activity assays are often not available for certain families of membrane proteins, e.g. ion channels and transporters, the successful overexpression is often not straightforward.

The use of the green fluorescent protein (GFP) as a folding indicator has revolutionized expression screening of membrane proteins in bacterial systems (Drew, von Heijne et al. 2001; Kawate and Gouaux 2006; Hammon, Palanivelu et al. 2009). When fused to the C-terminus of a membrane protein, GFP moiety folds properly and becomes fluorescent only if a membrane protein is also properly folded and stably inserted into the membrane (Drew, Lerch et al. 2006). C-terminal GFP fusions allow over-expressed membrane proteins to be monitored in intact cells by measuring in cell fluorescence (Hammon, Palanivelu et al. 2009). Since GFP does not unfold under standard SDS-PAGE conditions, membrane protein–GFP fusions can also be visualized directly by in-gel fluorescence (Drew, Lerch et al. 2006). In addition, the GFP moiety greatly facilitates purification and characterization of the membrane protein-GFP fusions (Kawate and Gouaux 2006).

While GFP fluorescence reveals the well folded population of a membrane protein, it does not quantify the total expression level i.e. the amount of produced protein that includes misfolded and aggregated protein. Recently, an elegant method was developed that allows detection of well folded and aggregated proteins simultaneously (Geertsma, Groeneveld et al. 2008). The method combines in-gel fluorescence and anti His-tag Western blot analysis to follow differential migration of folded and aggregated GFP fusion proteins during SDS-PAGE (Figure 2.1.4). Determination of the ratio of both species accelerates the optimization of the functional overexpression of membrane proteins.

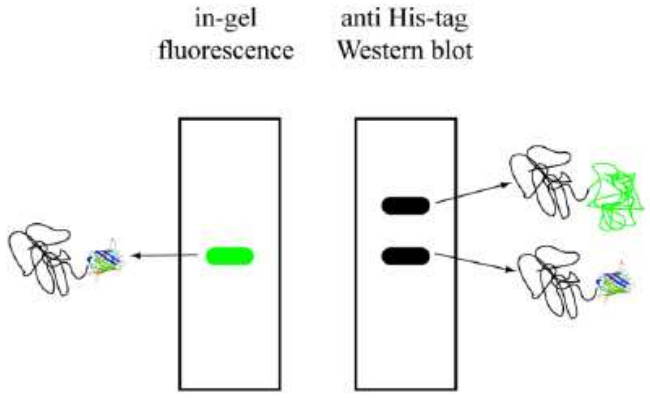



**Figure 2.1.4. GFP as expression and folding indicator**
Green fluorescent protein (GFP) can be used as a folding indicator. C-terminally fused GFP folds properly and shows fluorescence only if the preceding protein is also folded. In contrast, if the preceding protein is aggregated, GFP is misfolded and lacks fluorescence. Since GFP keeps its conformation during SDS-PAGE analysis, it can be used to directly trace the folded fraction of the target protein within the gel (left). Since folded and unfolded GFP fused to SDS denatured protein have different migration behaviors, they can be separated and thus used to analyze the ratio of folded (right, lower band) to aggregated protein (right, upper band) expression levels.

### 2.1.3.1 Whole cell fluorescence-based expression screening of prokaryotic ClC homologs

All prokaryotic ClC homologs were cloned as C-terminal fusions with GFP and $His_{10}$-tag (Figure 2.1.3) under the control of the $P_{BAD}$ promoter. The $P_{BAD}$ (araBAD) promoter was chosen because it offers an expression strategy that allows a tight regulation with dose-dependent induction, in that way combining gentle expression conditions with high expression yields. The inducer in this system is arabinose. In arabinose-free medium the AraC dimer binds two distant parts of the araBAD-operon and suppresses transcription through DNA loop formation. Upon arabinose binding, the AraC dimer rearranges and binds to adjacent regions in the araBAD-operon thus releasing the loop and activating transcription of the target gene (Schleif 2003).

For expression screening the ligation mixtures were transformed into the *E. coli* strain MC1061 that is incapable of metabolizing arabinose due to a disrupted *ara* gene.

In order to rapidly test for protein expression, all experiments were carried out in a high-throughput manner. The bacterial cultures were grown in 96 well plates with only 0.7 ml culture medium. This allowed a simultaneous screening of different homologs and inducer concentrations. The expression medium in these studies was Terrific Broth (TB) and the induction temperature was 25°C.

| colony | AfClC | CmClC | DaClC | HaClC | HvClC | HbClC | MsClC | MbClC | MmClC | MtClC | MaClC | Mb1ClC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2524 | 8742 | 3075 | 3552 | 8687 | 7632 | 5664 | 6597 | 10627 | 6022 | 8158 | 7335 |
| 2 | 2695 | 10170 | 3401 | 3424 | 24748 | 7149 | 5982 | 8175 | 8989 | 6786 | 1763 | 7869 |
| 3 | 2728 | 10733 | 3560 | 3843 | 10150 | 7356 | 5807 | 7899 | 10134 | 4538 | 6890 | 7364 |
| 4 | 2501 | 11390 | 4293 | 1479 | 10513 | 7710 | 4207 | 7509 | 9248 | 6416 | 7561 | 7714 |
| 1 | 4675 | 18962 | 2463 | 14489 | 11103 | 5314 | 10181 | 4983 | 4181 | 9980 | 11490 | 1649 |
| 2 | 4709 | 18834 | 2466 | 11950 | 10662 | 5460 | 11420 | 6124 | 4295 | 9858 | 11864 | 1614 |
| 3 | 4684 | 20047 | 2503 | 1576 | 10000 | 5850 | 10698 | 5588 | 4131 | 6153 | LacS | negative |
| 4 | 4864 | 21999 | 2742 | 14961 | 8636 | 6104 | 4801 | 5988 | 4246 | 9538 | positive | control |
| | Mb2ClC | MmzClC | PcClC | PlClC | PvClC | RpClC | ScClC | SfClC | TvClC | ClClC | control | |

**Figure 2.1.5. Whole cell fluorescence of 22 prokaryotic ClC homologs following induction with $10^{-2}$ % arabinose and overnight expression at 25°C.**
Tested ClC homologs can be divided into four groups based on the specific fluorescence intensity compared to the negative and positive control. Each homolog is represented with four independent colonies. Two homologs (yellow) show very low fluorescence, almost close to the background. Eight homologs (olive green) show low to intermediate fluorescence. Six homologs (sea green) show intermediate to high fluorescence. Six homologs (fluorescent green) show high to very high fluorescence and can be considered promising based on this experiment. Only three colonies (red) out of 88 do not contain an insert which illustrates the high efficiency of FX cloning. As a positive control *E. coli* transformed with the plasmid containing LacS protein from *Streptococcus termophilus* fused to GFP was used. As a negative control *E. coli* cells transformed with a vector without GFP were used.

The transformed bacterial cells were grown at 37°C to an $OD_{600}$=0.5, cooled down to 25°C and induced with a range of arabinose concentrations. After overnight expression, the cells were pelleted, resuspended in PBS buffer and analyzed for fluorescence. The measured fluorescence is normalized by the total amount of bacteria and listed as specific fluorescence. Figure 2.1.5 presents a 96 well plate with 22 ClC homologs tested for whole cell fluorescence after induction with $10^{-2}$ % arabinose.

When comparing specific fluorescence values of the 22 homologs tested in this experiment with the background fluorescence of the negative control, it can be concluded that all homologs express folded fusion proteins at different levels. The amounts of fluorescence vary up to ten fold between different homologs. Based on the specific fluorescence, all investigated proteins are classified into four groups. The first group contains two proteins, AfClC and PcClC**,** with very low fluorescence levels, less than two times over the background. The second group encompasses eight ClC homologs (DaClC, HaClC, MsClC, MtClC, Mb2ClC, RpClC, SfClC and TvClC) with low to intermediate fluorescence, two to four times higher than the background. In the third group we find six ClC homologs (HbClC, MbClC, MmClC, MaClC, Mb1ClC, ClClC) showing intermediate to high fluorescence levels, four to six times over the background. The fourth group is the most promising one with six proteins (CmClC, HvClC, MmzClC, PlClC, PvClC, and ScClC) that show high to very high fluorescence, in some cases more than ten times over the background. It is important to emphasize that the described classification is valid for the presented experimental condition. The induction with $10^{-2}$ % arabinose was chosen, since for all investigated concentrations it illustrates the differences in fluorescence between homologs best. Though, this inducer concentration is not optimal for all proteins to yield the highest amount of well folded fusion proteins.

Whole cell fluorescence can be used to estimate the optimal inducer concentration for a given membrane protein-GFP fusion. The expression of all 27 cloned ClC homologs was investigated for four arabinose concentrations. Figure 2.1.6 shows the specific fluorescence of three ClC homologs at specified inducer concentrations. The increased arabinose concentration has a different influence on the amount of the fluorescent fusions in the presented proteins. For all three fusion proteins the lowest arabinose concentration ($10^{-4}$ %) does not induce significant expression, as the fluorescence values are close to those of the negative control. Interestingly, the higher arabinose concentrations do not give similar results for all the homologs. In the case of the Rm1ClC homolog, it appears that increased inducer concentration does not have a negative effect on the yield of functional protein. The fluorescence is slowly and steadily increasing and all three expression conditions provide high amounts of properly folded fusion protein. For the ClClC homolog, $10^{-2}$ % arabinose in the medium induces the highest amount of fluorescent protein and can be considered optimal based on this experiment. For EcClC, the well

characterized homolog from *E. coli*, the optimal inducer concentration is $10^{-3}$ %, which is in agreement with previously published results (Geertsma, Groeneveld et al. 2008).

| | Rm1ClC | | | | ClClC | | | | EcClC | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [Ara] | $10^{-4}$% | $10^{-3}$% | $10^{-2}$% | $10^{-1}$% | $10^{-4}$% | $10^{-3}$% | $10^{-2}$% | $10^{-1}$% | $10^{-4}$% | $10^{-3}$% | $10^{-2}$% | $10^{-1}$% |
| colony 1 | 2110 | 11898 | 15440 | 15247 | 1851 | 7555 | 10225 | 6542 | 2659 | 18695 | 11288 | 6049 |
| 2 | 1949 | 11161 | 15698 | 17316 | 2031 | 7574 | 9926 | 5826 | | | | |
| 3 | 1989 | 12266 | 15877 | 17948 | 1756 | 4572 | 9009 | 5180 | Negative control | | | |
| 4 | 2072 | 11713 | 15577 | 17658 | 1967 | 7787 | 10037 | 6819 | 1901 | 1812 | 1735 | 1432 |

**Figure 2.1.6. Whole cell fluorescence of Rm1ClC, ClClC and EcClC at different inducer concentrations.**
Whole cell fluorescence can be used to estimate the optimal inducer concentration that gives the highest amount of well folded membrane protein. In the case of Rm1ClC, increasing arabinose concentrations lead to higher fluorescence. In the case of ClClC, an arabinose concentration of $10^{-2}$ % is optimal, since it gives the highest amount of fluorescent protein. The optimal inducer concentration for EcClC is $10^{-3}$ % arabinose.

Expression screening of prokaryotic ClC homologs by measuring whole cell fluorescence has provided initial information on the expression levels and optimal inducer concentrations for each membrane protein. A set of candidate proteins with promising overexpression potential were identified in that way. Still, this simple GFP screen did not provide any information on the amount of misfolded membrane proteins produced. Measured fluorescence also did not reveal the integrity of the membrane protein-GFP fusion as degradation of the expressed target protein could also contribute to the fluorescence. In order to study these properties, the well expressed candidates were subsequently analyzed by in-gel fluorescence and anti-His Western blot.

### 2.1.3.2 Expression screening of prokaryotic ClC homologs by in-gel fluorescence and Western blot

After quantifying whole cell fluorescence, the same cultures were used to prepare samples for SDS-PAGE. For comparison, similar quantities of cells were analyzed in order to correlate the specific amounts of overexpressed proteins. Following the electrophoretic separation of proteins, the fluorescence in the SDS-PAGE gel was analyzed. The size of the target protein-GFP fusion was estimated from the position of the fluorescent band on the gel. In that way also degradation products were detected. The same gel was subsequently analyzed by Western blotting on a PVDF membrane. Due to the presence of a $His_{10}$-tag at the C-terminus, the fusion proteins can be detected with appropriate

antibodies. The combination of in-gel fluorescence and anti His-tag Western blot analysis of the same gel provides complementary information on the amount of expressed proteins as described in Figure 2.1.4.

For ClC homologs that showed low to intermediate expression, the in gel fluorescence signals correlate well with the fluorescence observed in whole cells (Figure 2.1.7). All presented ClC-GFP fusions migrate at the expected molecular weight confirming the integrity of the membrane protein part. In the case of the homolog TvClC a strong fluorescent degradation band is detected. In almost all samples only one band corresponding to the folded protein is detected by anti His-tag immunoblot. Only for TvClC a strong upper band of aggregated protein was identified.

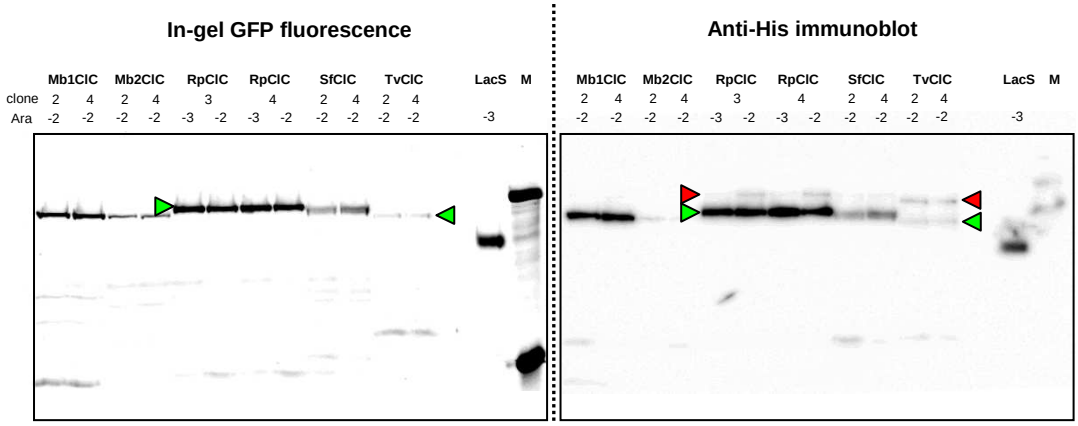



**Figure 2.1.7. Analysis of expression using GFP folding assay.**
Five ClC homologs previously classified to have low to intermediate expression level were analyzed by in-gel fluorescence (left panel) and anti-His immunoblot (right panel). Two clones for each homolog were tested. Arabinose concentrations used for induction are indicated (-3 and -2 for $10^{-3}$ % and $10^{-2}$ %, respectively). The upper fluorescent band in the marker corresponds to 75 kDa, the lower one to 25 kDa. Green triangles indicate well folded fluorescent protein, whereas red triangles indicate misfolded non-fluorescent protein. Both analyses were performed using the same SDS-PAGE gel.

Figure 2.1.8 shows the analysis of two ClC homologs from the same bacterium *Ralstonia metallidurans*, Rm1ClC and Rm2ClC. Already whole cell fluorescence experiments revealed substantial difference in expression levels between the two proteins, which is additionally confirmed by in-gel fluorescence. Since Rm1ClC was identified as one of the proteins with the highest overexpression potential, the quality control with anti His-tag Western blot analysis was of great importance. Only a negligible amount of aggregated fusion protein was detected even at the highest inducer concentration. These results make Rm1ClC very attractive for more detailed examinations.

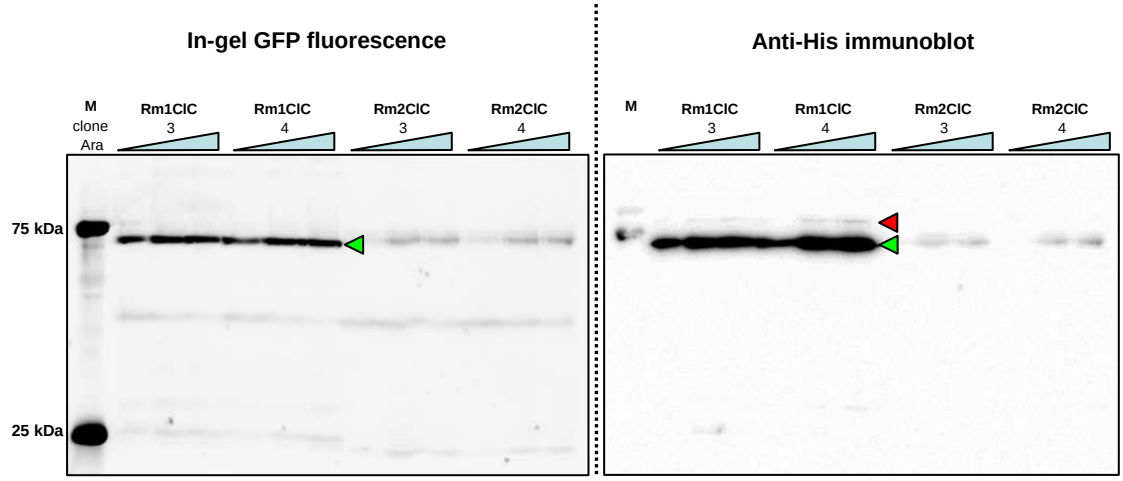



**Figure 2.1.8. Analysis of the *Ralstonia metallidurans* ClC homologs by GFP folding assay.**
Rm1ClC and Rm2ClC were analyzed by in-gel fluorescence (left panel) and anti-His immunoblot (right panel). Two clones for each homolog were tested. Increasing arabinose concentration of $10^{-3}$, $10^{-2}$, $10^{-1}$ % (w/v) was used for induction (blue triangle). Green triangles indicate well folded fluorescent protein, whereas red triangles indicate misfolded non-fluorescent protein. Both analyses were performed with the same SDS-PAGE gel.

A few ClC homologs were classified as promising based on their whole cell fluorescence experiments, but did not show satisfying behavior in continuing assays. For example, although the HvClC homolog is expressed at high levels in the folded form, the amount of aggregated protein is also high (Figure 2.1.9). At all tested inducer concentrations, folded and misfolded protein are present in the similar amounts. This could indicate that the expression of HvClC does not overload the folding machinery of the cell, but rather that this membrane protein is lacking a specific factor from its natural environment.

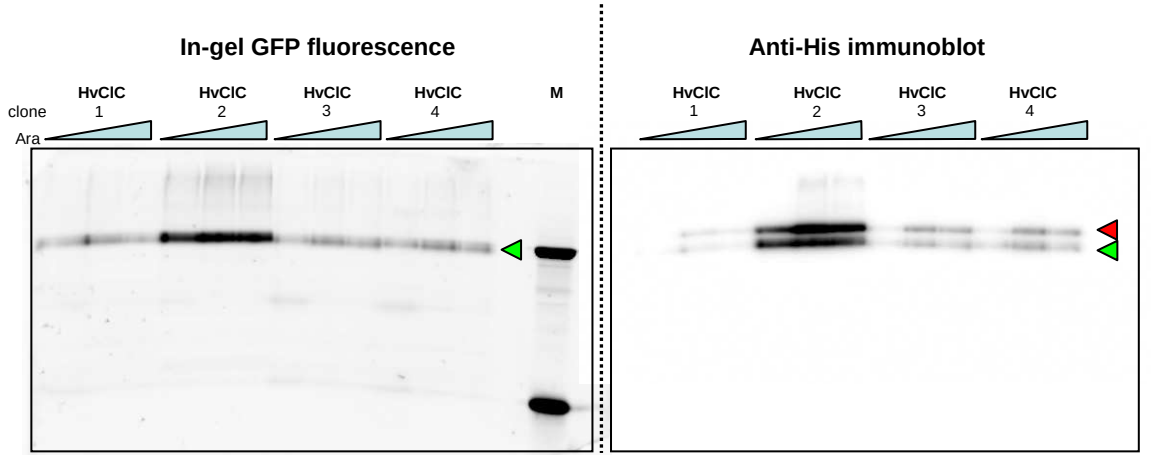



**Figure 2.1.9. Analysis of the *Haloferax volcanii* ClC homolog by GFP folding assay.**
HvClC was analyzed by in-gel fluorescence (left panel) and anti-His immunoblot (right panel). Four independent colonies were tested for expression. Colony No. 2 shows a much higher expression level than the other three colonies for unknown reasons. Increasing arabinose concentration of $10^{-3}$, $10^{-2}$, $10^{-1}$ % (w/v) was used for induction (blue triangle). Green triangles indicate well folded fluorescent protein, whereas red triangles indicate misfolded non-fluorescent protein. Both analyses were performed using the same SDS-PAGE gel.

An interesting behavior was discovered for ScClC, where aggregated protein dominates at arabinose concentrations of $10^{-2}$ and $10^{-1}$ % (Figure 2.1.10). In the case of MmzClC homolog there is no correlation between the measured whole cell fluorescence values and the results of in-gel fluorescence. The reason for this observation might be membrane protein degradation after expression, which leaves the folded GFP moiety intact.

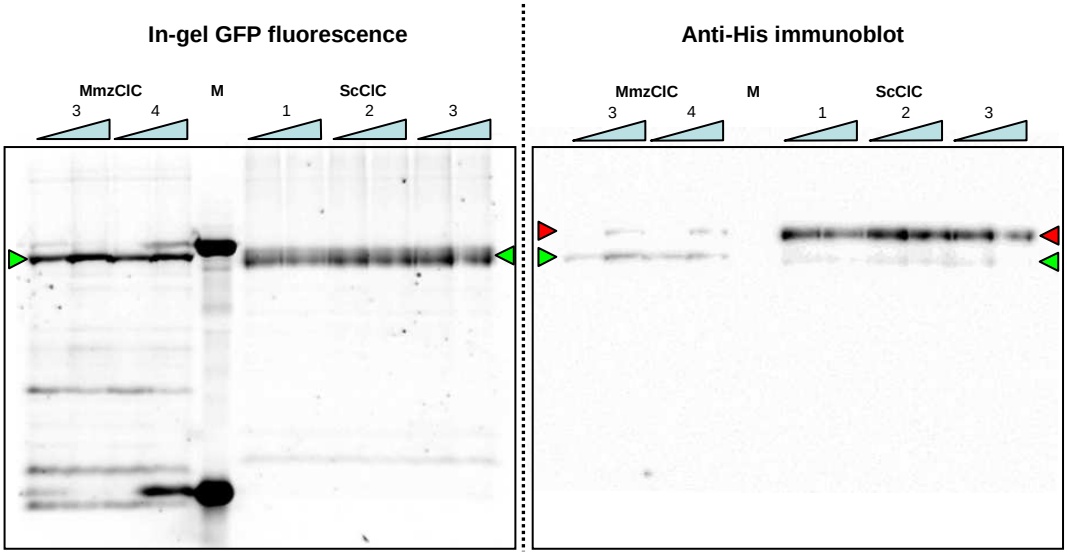



**Figure 2.1.10. Analysis of expression using GFP folding assay.**
MmzClC (represented with two clones) and ScClC (represented with three clones) were analyzed by in-gel fluorescence (left panel) and anti-His immunoblot (right panel). Arabinose concentration used for induction was $10^{-2}$ and $10^{-1}$ % (blue triangles). The upper fluorescent band in the marker corresponds to 75 kDa, the lower one to 25 kDa. Green triangles indicate well folded fluorescent protein, whereas red triangles indicate misfolded non-fluorescent protein. In the samples of MmzClC fluorescent degradation bands are noticeable. In the samples of ScClC the aggregated protein is expressed at high amounts. Both analyses were performed using the same SDS-PAGE gel.

The extensive analysis of expression properties followed by quality control experiments have allowed for the classification of cloned prokaryotic ClC homologs and the selection of promising candidates for closer examination. Table 2.1.2 summarizes those results. Only one homolog, Rm1ClC, showed very high expression level, comparable with the expression level of EcClC that is overexpressed in its native host. Nine additional ClC proteins were selected for further characterization. Overall, the tendency for aggregation for this set of prokaryotic ClC proteins was low. Still, the optimal temperature and inducer concentration will have to be identified for each homolog.

Though it is interesting to correlate the microbial origin of the homolog with its expression behavior, it is difficult to draw conclusions. The homologs from thermophilic archaea and δ/ε proteobacteria (labeled blue in Table 2.1.2 and Figure 2.1B) generally express at low level and some of them show aggregation and degradation, combining in

that way many undesirable properties. From other four “phylogentic” groups at least two homologs showed satisfactory expression levels and were chosen for further studies.

| ClC homolog | Expression level<br><br>Whole cell GFP fluorescence | Aggregation (upper) band<br><br>Western blot | Degradation bands<br>In gel GFP fluorescence | Selected for further examination |
|---|---|---|---|---|
| AfClC | very low | - | - | |
| MtClC | low | + | + | |
| PtClC | very low | - | - | |
| TvClC | low | + | + | |
| MsClC | low | + | + | |
| CmClC | intermediate | - | + | |
| HyClC | low | - | - | |
| MmClC | intermediate | + | - | |
| HbClC | intermediate | - | + | |
| DaClC | very low | - | - | |
| PcClC | intermediate | - | + | |
| SfClC | low | - | - | |
| ClClC | intermediate | - | - | + |
| CtClC | intermediate | - | - | |
| PlClC | high | - | + | + |
| PvClC | high | - | - | + |
| MaClC | intermediate | - | ++ | |
| MmzClC | high | - | ++ | + |
| Mb2ClC | very low | - | - | |
| Mb1ClC | intermediate | - | - | + |
| MbClC | intermediate | + | - | + |
| HaClC | low | + | - | |
| HvClC | high | ++ | - | + |
| ScClC | high | ++ | - | + |
| Rm1ClC | very high | - | - | + |
| Rm2ClC | low | - | - | |
| RpClC | intermediate | - | - | + |

**Table 2.1.2. Summary of expression screening of prokaryotic ClC homologs.**
The 27 ClC homologs are listed as in the phylogram obtained from a sequence alignment (Figure 2.1B/C). The expression levels determined by whole cell fluorescence measurements are shown. The presence of a prominent aggregation band on the Western blot is indicated as ++, moderate as + and no aggregation as - . Degradation bands visible on in-gel fluorescence are labeled as ++ when strong, + when moderate and – when absent. Homologs selected for further examination are indicated (+)

### 2.1.4 Extraction tests on selected membrane proteins

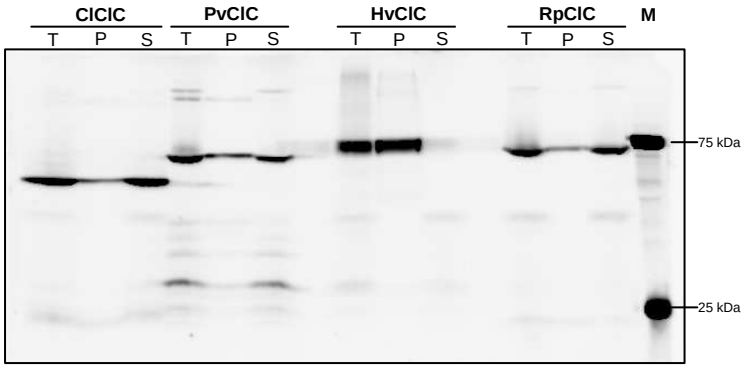
For purification purposes, membrane proteins need to be solubilized. For this purpose, different types of detergents are used. The role of detergent is to keep the target protein in its functional, folded state in the absence of a membrane by mimicking the natural hydrophobic environment of the lipid bilayer (Prive 2007). The identification of a suited detergent for a particular membrane protein is a challenging step. Apart from being able to extract the protein from the membrane, the detergent must stabilize the protein in solution during the course of experiment. So far, the most successful detergents for the crystallization of membrane proteins have been nonionic detergents (mild detergents) including glucosides, maltosides, polyoxyethylenes and the zwitterionic amine oxides.

For initial extraction tests of ClC homologs from the membrane the mild detergent dodecyl-β-D-maltoside (DDM) was chosen. The experiments were performed on a small scale level from whole cell lysate without prior isolation of the membrane fraction. After expression of the protein, the cells were broken and the membrane proteins were solubilized by addition of 1 % DDM (w/v) and incubation for 1 h on ice. To remove the insoluble cell debris and aggregated (unsolubilized) membrane proteins, the samples were centrifuged at high velocity. The amount of extracted membrane proteins in the supernatant was subsequently compared with the total amount of expressed protein before centrifugation.

The expression of ClC proteins fused to GFP has been advantageous also in these analyses. The fusion proteins have been investigated by in-gel fluorescence and by measuring GFP fluorescence in solution. Both procedures allowed quantification of fluorescent material that can be extracted from the membrane.

Figure 2.1.11 illustrates extraction tests on four ClC homologs by in-gel fluorescence. For each homolog three samples were analyzed: total material before centrifugation (T), pellet - unsolubilized fraction after centrifugation (P) and supernatant - solubilized fraction after centrifugation (S). It is evident from the gel (Figure 2.1.11A) that the proteins ClClC and RpClC can be almost completely extracted from the membrane, as the fluorescent band in the pelleted sample is minor. The amount of fluorescent protein in the pellet of PvClC is higher than in the previous two homologs. In the case of HvClC, the fluorescent protein in the solubilized fraction is almost unnoticeable, which might mean that the protein is either not extractable or it rapidly precipitates after extraction. This result is not so surprising when considering the behavior of HvClC in the previously performed quality control assays. Figure 2.1.11B shows the quantification of solubilized fraction for each fusion protein applying densitometric analysis of the fluorescent bands.

A)




B)

| ClClC | | PvClC | | HvClC | | RpClC | | |
|---|---|---|---|---|---|---|---|---|
| T | S | T | S | T | S | T | S | |
| 490000 | 480000 | 570000 | 280000 | 550000 | 65000 | 370000 | 360000 | measured fluorescence |
| 100% | 98% | 100% | 49% | 100% | 12% | 100% | 97% | solubilized fraction |

**Figure 2.1.11. Extraction of ClC homologs from the membrane with dodecyl-β-D-maltoside (DDM).**
After expression and cell lysis, samples were incubated with 1 % DDM followed by high-spin centrifugation. Total material before centrifugation (T), pellet (P) and supernatant (S) after centrifugation were analyzed by in-gel fluorescence.

A) SDS-PAGE gel with fluorescent fusion proteins. The proteins ClClC, PvClC and RpClC are extractable with DDM as the fluorescent band in the solubilized fraction has a strong intensity compared to the starting sample. HvClC is not extractable from the membrane as most of the fusion protein is found in the pelleted sample.

B) Quantification of extractability by densitometric analysis of the fluorescent bands. ClClC and RpClC are almost completely extracted from the membrane with DDM.

Out of ten ClC homologs selected for closer examination, after expression tests six were extractable with DDM at a satisfying level (minimum 50 %). For the remaining four homologs HvClC, MbClC, Mb1ClC and ScClC the amount of extractable protein was too low to pursue with further experiments that are part of the precrystallization screening. It is necessary to mention that a more thorough expression optimization could have been performed in order to express these proteins in an improved condition, which could have led to a better detergent extractability. However, I limited the more extensive investigations only to the six homologs that were extractable with DDM at the first place: ClClC, MmzClC, PlClC, PvClC, Rm1ClC and RpClC.

### 2.1.5 Test purification of prokaryotic ClC homologs

In the standard purification procedure of membrane proteins for crystallization, two purification steps after extraction from the membrane are commonly used. The first step is immobilized metal affinity chromatography (IMAC) followed by size exclusion chromatography (SEC).

IMAC is a separation technique that is based on the coordination between imidazole groups and metal ions. Recombinant proteins fused with a histidine tag ($His_6$ or $His_{10}$-tag) bind to immobilized divalent ions like $Ni^{2+}$, $Co^{2+}$, $Fe^{2+}$ and $Zn^{2+}$. In this way target protein can be retained on $Me^{2+}$ chelating matrix and purified from mixture of proteins in the cell lysate or solubilized membrane fraction.

SEC separates proteins based on their size and shape and therefore enables separation from impurities with very different molecular weight from the target protein. Additionally and more importantly, SEC serves as one of the most useful tools for separating high molecular aggregates and monitoring the monodispersity, stability and oligomeric state of a target protein. A monodisperse and folded protein generally gives a single symmetrical peak, while a polydisperse, unstable or unfolded protein typically gives multiple asymmetric peaks (Barth, Boyes et al. 1994). A monodisperse size exclusion profile is a good indicator of a membrane protein that is of suitable quality for further biochemical and crystallization studies (Kawate and Gouaux 2006). Testing many protein constructs for monodispersity in the precrystallization screening is thus the most reliable criterion for the selection of promising candidates.

The six selected ClC homologs were tested for monodispersity in the following way: Each ClC protein-GFP fusion was expressed in 500 ml TB medium using the conditions determined in the expression screening. The cells were harvested and broken and the membrane fraction was isolated by high-spin centrifugation. The membrane proteins were extracted from the membrane with DDM and the unsolubilized fraction was pelleted by high-spin centrifugation. The supernatant after centrifugation was used for protein purification with IMAC. In this experiment Ni-NTA agarose (Qiagen) was used for batch binding of fusion proteins. The affinity matrix with immobilized protein was washed with buffer containing low concentration of imidazole to remove non-specifically bound proteins. Subsequently, the bound fusion protein was eluted from the resin with the higher imidazole concentration. After IMAC purification, the GFP fusion proteins were subjected to size exclusion chromatography on a Superdex 200 column (S200 GE Healthcare). The calibration of the column with known protein standards provides the relationship between protein molecular weight and the retention volume. In the case of membrane proteins that are surrounded by a detergent micelle, though, the apparent molecular weight is higher than expected because of the contribution of the detergent.

Figure 2.1.12 shows size exclusion profiles for the homologs Rm1ClC, RpClC and ClClC on S200. The protein was detected by measuring its absorbance at 280 nm. The lower histograms show the measured fluorescence in the fractions eluted from the S200 column.

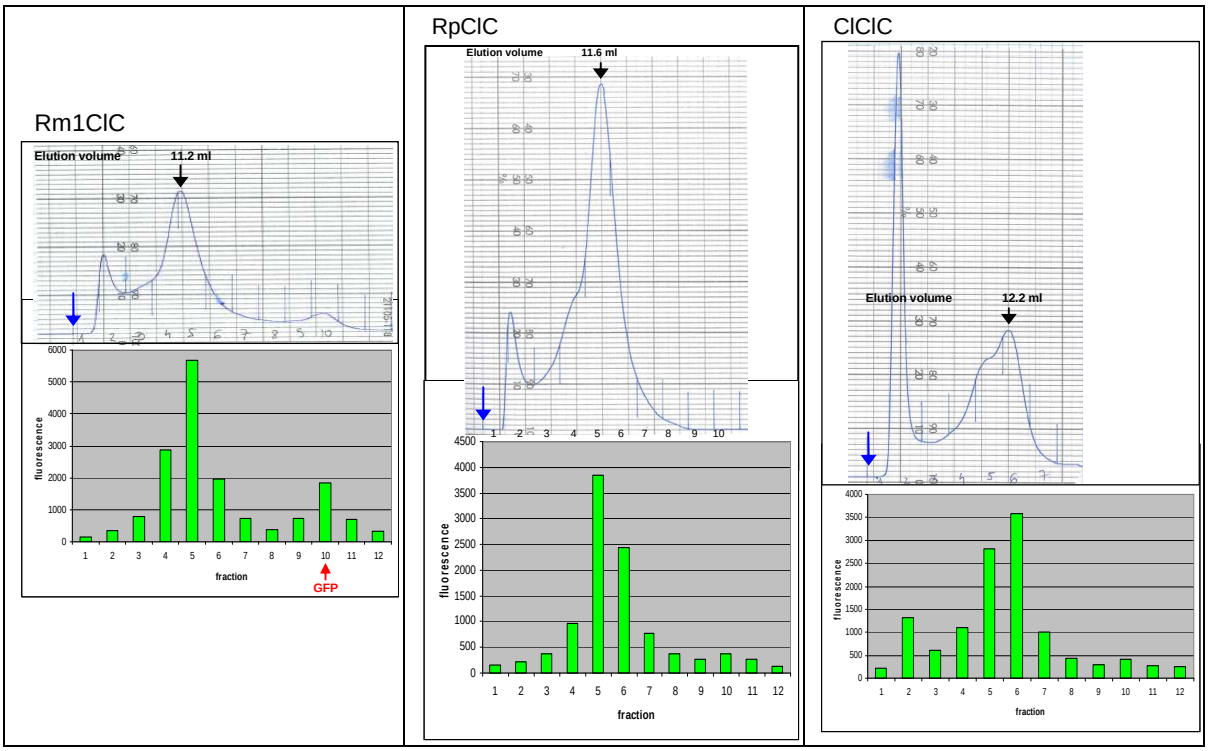



**Figure 2.1.12. Size exclusion chromatography (SEC) of Rm1ClC, RpClC and ClClC in DDM.**
Purified GFP fusions of ClC homologs were analyzed by SEC on S200 column. The chromatograms present absorbance at 280 nm during the separation, the histograms present fluorescence values in each eluted fraction. The blue arrow is indicating the void volume (7 ml). For all three homologs the fluorescent proteins have elution volume that corresponds to well folded proteins. In the case of Rm1ClC and RpClC the elution peaks are monodisperse. In the case of ClClC, there are higher molecular aggregates at the void volume and the elution peak contains two populations with different oligomeric state.

The elution profiles for all three homologs looked very encouraging. Rm1ClC-GFP shows a symmetric peak with an elution volume of 11.2 ml, which corresponds to a molecular weight of a dimeric protein. Based on the absorbance in the second fraction, it appears that there is a certain amount of aggregated protein that elutes at the void volume of the column, but as the fluorescence in that fraction is negligible, the absorbance is probably due to impurities.

RpClC-GFP elutes at 11.6 ml, corresponding to the approximate molecular weight of a dimer. The difference in the elution volumes between the Rm1ClC-GFP and RpClC-GFP fusions, that have a similar molecular weight, might result from slight differences in

the shape of the two molecules. The elution peak of RpClC contains a second smaller peak at lower elution volumes (fraction four).

The size exclusion profile of ClClC indicates that the solubilized protein is not monodisperse. While the largest amount of the fluorescent protein is eluted at 12.2 ml, there is a second peak eluting at a lower volume, which indicates the presence of a higher oligomeric state. The two populations could represent the dimeric and a monomeric species of the protein, based on the estimate of the elution volumes. In addition, higher molecular aggregates of ClClC eluting at the void volume were found.

Figure 2.1.13 represents size exclusion profiles for the homologs MmzClC, PlClC and PvClC. It is apparent that none of the three proteins is monodisperse. High amounts of aggregated proteins eluting in the void volume and degradation products in the form of the free GFP were detected by following the absorbance at 280 nm and by measuring the fluorescence of the collected fractions. MmzClC shows a severe degradation that was already observed during expression studies by in gel fluorescence. The fluorescence values for this protein are very low comparing to the other tested homologs. It is possible that the fusion protein is degrading during purification leaving the soluble GFP intact, and causing the membrane protein to aggregate.

The similar behavior of PlClC and PvClC might be correlated with the microbial origin of these two proteins, which come from closely related bacteria (Figure 2.1.1B). The highest amounts of both proteins were found as high molecular aggregates. However, there is a small amount that elutes at 11.6 ml and 11.8 ml, respectively, that might correspond to properly folded protein. It is possible that additional expression optimization and the choice of a different detergent could improve the biochemical behavior of PlClC and PvClC, but those experiments have not been performed during the course of this project.

In summary, three ClC homologs (MmzClC, PlClC and PvClC) were characterized as polydisperse and not suitable for further biochemical and crystallization studies. The proteins ClClC and RpClC required additional examinations to be properly classified. The Rm1ClC homolog showed the most promising behavior in the performed test for monodispersity in detergent solution.

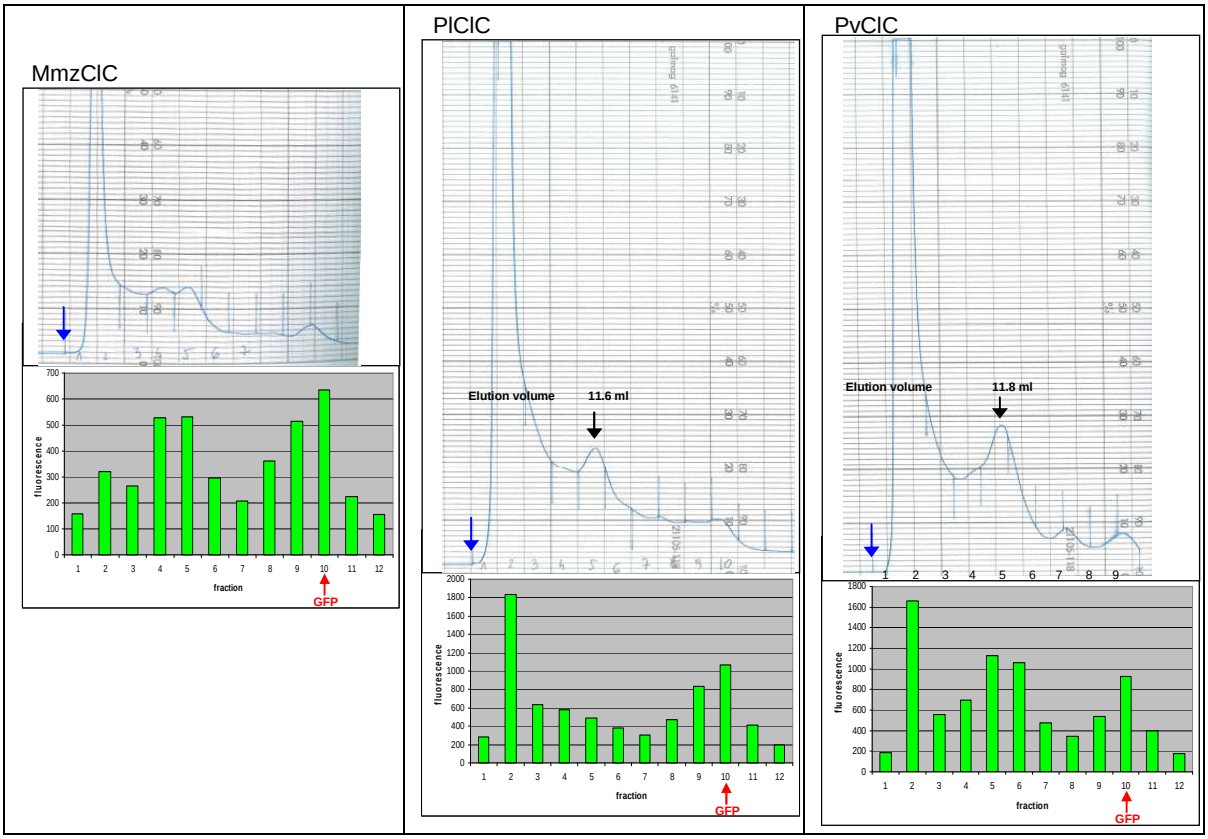



**Figure 2.1.13. Size exclusion chromatography (SEC) of MmzClC, PlClC and PvClC in DDM.**
Purified GFP fusions of ClC homologs were analyzed by SEC on S200 column. The chromatograms present absorbance at 280 nm during the separation, the histograms present fluorescence values in each eluted fraction. The blue arrow is indicating the void volume (7 ml). All three homologs are polydisperse and their chromatograms show strong aggregation peaks. In the case of MmzClC the fluorescence due to degradation to free GFP is higher than from the fusion. In the case of PlClC and PvClC, most of the protein is in the aggregated state and elutes in the void volume.

## 2.1.6 Further biochemical analyses of RpClC and ClClC

After testing for monodispersity by size exclusion chromatography, three ClC homologs were selected as promising and subjected to more extensive examinations. In this chapter, follow-up experiments on RpClC and ClClC are presented.

The expression and purification of RpClC-GFP fusion was carried as described before, starting from the culture volume of 1 l. Following IMAC purification, the eluted and concentrated protein was injected onto a S200 column equilibrated with buffer containing DDM (Figure 2.1.14). Interestingly, the size exclusion profile is very similar to the one already observed (Figure 2.1.12A). A "shoulder" in the fourth fraction is preceding the main elution peak coming at the volume of 11.7 ml. The protein from the peak fraction (No. 5) was reinjected on S200 after 24 h incubation at 4ºC (Figure 2.1.12B). The elution profile is changed, indicating that RpClC is not stable under the applied conditions. Two

protein populations can be observed, one eluting at the volume corresponding to the "shoulder" in the first SEC (10.2 ml), a second one eluting at the volume of the main peak in the first SEC (11.7 ml). The height of the 10.2 ml peak suggests that the protein has a tendency to form higher oligomers over time.

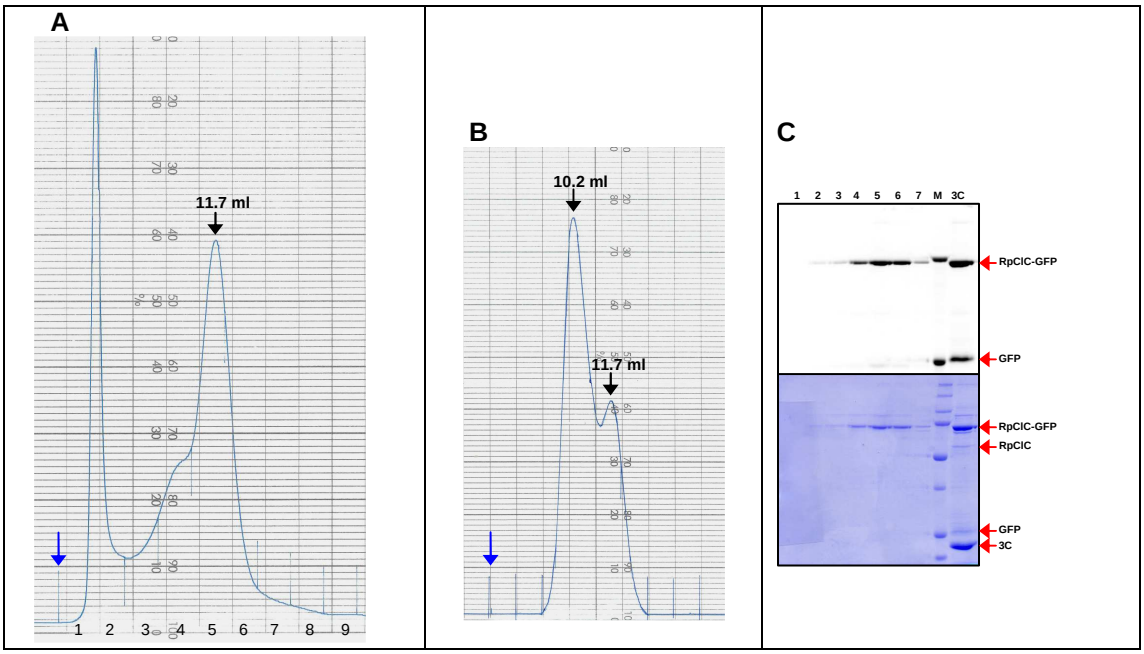



**Figure 2.1.14. Test purification of RpClC-GFP fusion in DDM.**
A) Size exclusion chromatography of RpClC-GFP fusion purified with IMAC. The protein elutes at 11.7 ml (S200 column). The elution peak is not symmetrical due to a shoulder in the fourth fraction. The blue arrow indicates the void volume (7 ml).
B) Size exclusion profile of RpClC-GFP eluted in fraction 5 (A) after reinjection onto S200 after 24 h. The protein shows two peaks.
C) SDS-PAGE analysis (top - in gel fluorescence, bottom - coomassie stained gel) of the SEC fractions 1-7 from figure A. The sample labeled 3C presents RpClC-GFP fusion incubated for 3 h with 3C protease. The protease cleavage turned out not to be complete.

For crystallization experiments the GFP-$His_{10}$ part needs to be removed from the target protein. Therefore, RpClC-GFP fusion was digested by 3C protease, which cleaves at a specific site at the C-terminus of the membrane protein (Figure 2.1.12B, lane 3C). The 3C protease is very efficient in cutting of the His-tag from both soluble and membrane proteins at different buffer compositions and temperatures. However, in the case of RpClC fusion, the cutting efficiency was very low. The reason for this might be the inaccessibility of the cleavage site for the protease due to a special folding of the protein. This finding together with the results from size exclusion chromatography suggested that RpClC was not suitable for crystallization.

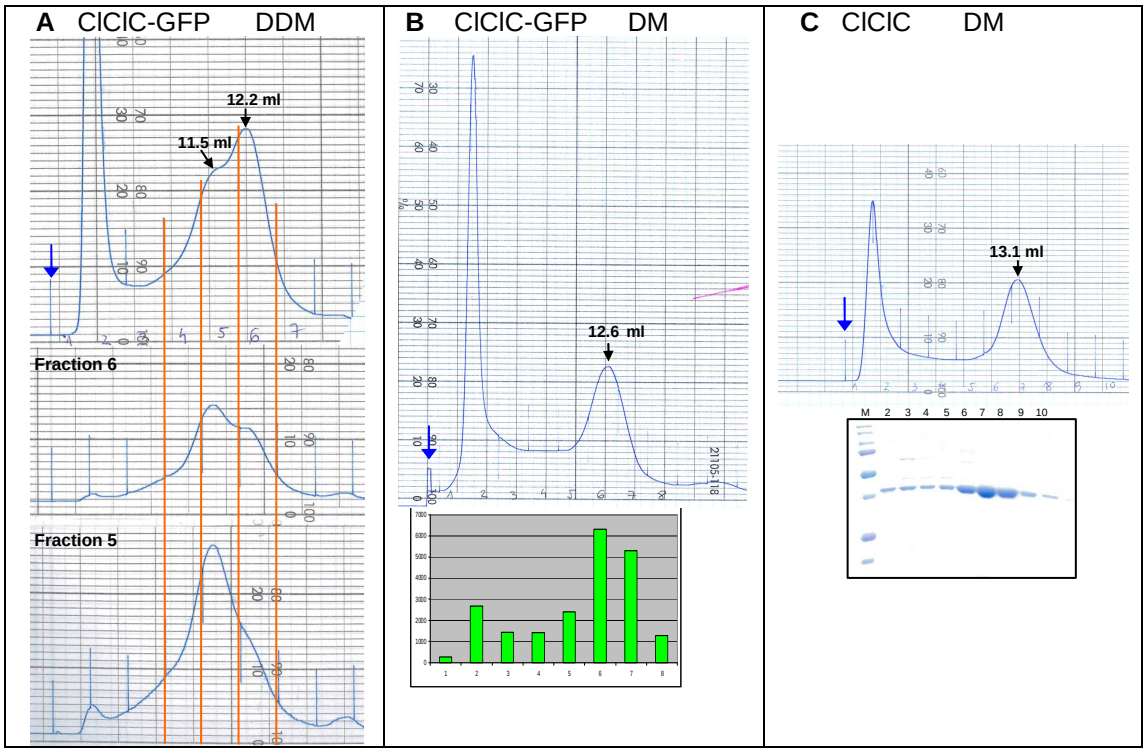



**Figure 2.1.15. Test purifications of ClClC in DDM and DM.**
A) Size exclusion chromatography of ClClC-GFP fusion in DDM. The protein is polydisperse with two populations eluting at 11.5 ml and 12.2 ml (S200 column). The blue arrow indicates the void volume (7 ml). After reinjection of fractions 5 and 6 independently on S200, the dominating component has the lower elution volume corresponding to the higher molecular weight.
B) Size exclusion profile of ClClC-GFP fusion in DM. The fluorescent fusion elutes at the volume of 12.6 ml. The protein is not monodisperse, because of the presence of aggregated protein eluting at the void volume.
C) Size exclusion profile of ClClC protein purified with IMAC and cleaved with the 3C protease to remove GFP-$His_{10}$ part (top). The ClClC protein elutes at 13.1 ml. The SDS-PAGE (bottom) of SEC fractions shows that the highest amount of the protein is eluted in the peak fractions.

The ClClC-GFP fusion has shown a polydisperse distribution in the previous SEC analysis. Besides the presence of aggregated protein, the protein eluted at higher volumes also appeared to exist in different oligomeric states, most probably in a monomer dimer equilibrium. After reinjection of fractions containing one of the two species on a S200 column, the fraction with the higher molecular weight was dominating (Figure 2.1.15A). In contrast to the elution in DDM, the size exclusion profile of ClClC-GFP in β-D-decyl-maltoside (DM) looks different (Figure 2.1.15B). Besides a certain amount of aggregated protein, a peak at 12.6 ml (corresponding to the higher elution volume in DDM) appears to be monodisperse.

The cleavage of the GFP from the ClClC-GFP fusion protein after IMAC purification was successful and the cleaved membrane protein was analyzed by SEC in DM (Figure 2.1.15C). The elution profile is very similar to the profile of the fusion protein, except for a shift in the elution volume to 13.1 ml due to the lack of GFP moiety. This

elution volume probably corresponds to a monomeric protein, which is not the native dimeric state for ClC proteins. It is conceivable that the harsher detergent DM shifts the equilibrium towards the monomeric state, while in DDM both oligomeric states are coexisiting. More sophisticated techniques such as multi angle static light scattering or analytical ultracentrifugation are necessary to determine the oligomeric state of ClClC with certainty. It might also be possible to stabilize the dimer by different detergents or by the addition of lipids during purification and thus make the protein a more promising candidate for crystallization screening.

### 2.1.7 Summary

To identify promising candidates for structural analysis 27 prokaryotic homologs of ClC chloride channels and transporters were cloned as C-terminal GFP fusion proteins under the control of pBAD promoter. Expression tests were performed by whole cell fluorescence measurements. Additionally, a combination of in-gel fluorescence and Western blot analysis allowed to judge the quality of expressed proteins at an early stage. Based on the expression level, ten ClC homologs were selected for closer examination. In extraction tests with DDM, six proteins showed a satisfying behavior and were further analyzed by size exclusion chromatography. During this procedure, three proteins (MmzClC, PlClC and PvClC) showed severe aggregation, which made them unsuitable for crystallization screening. Two proteins, ClClC and RpClC, showed initially promising SEC profile, but closer examination revealed that they are not stable in a set of chosen detergents. Overall, only one protein, Rm1ClC, was identified as highly expressed, monodisperse and stable. The work on Rm1ClC is presented in the following chapters.

### 2.1.8 Material and methods

**Preparation of genomic DNA (gDNA).** The genomic DNA was isolated from the dried cell pellets using a GenElute$^{TM}$ Bacterial Genomic DNA Kit (Sigma). Briefly, bacteria were lysed with lysozyme in a chaotropic salt-containing solution. Proteins were removed from DNA by digestion with proteinase K and incubation at 55ºC. DNA is bound to the silica-based membrane and the remaining lysate is removed by centrifugation. After washing to remove contaminants, the DNA is eluted with buffer into a collection tube.

**PCR.** For amplification of ClC genes different polymerases were used to obtain a full set of amplified products. The reaction mixtures were adjusted according to the

accompanying manuals for Phusion polymerase (FINNZYMES), KOD polymerase (Novagen) and thermoprime Mastermix (ABgene). Reactions were supplemented with 20 - 50 µg of genomic DNA, 25 µM forward and reverse primer and 3% DMSO. To avoid amplification of nonspecific sequence a touchdown PCR protocol was applied. The starting annealing temperature of 58.5 ºC was decreased in each of the first 15 cycles for 0.5 degree. In the following 15 cycles the annealing temperature was kept constant at 50ºC. Denaturation and elongation temperatures were chosen according to the protocol of the polymerase. The PCR was finished by a final elongation step for 5 minutes at the optimal polymerase temperature. The PCR products were separated by agarose gel electrophoresis and the expected DNA bands excised and purified with High Pure PCR Product Purification Kit (Roche).

**Ligation.** For ligation, the linearized plasmid was mixed with a digested insert at a molar ratio of 1:5, supplemented with reaction buffer and 400 U of T4 DNA ligase (NEB), adjusted to a final volume with water and incubated for 1 h at room temperature.

**Transformation of *E. coli*.** 60 µl of chemically competent *E. coli* cells MC1061 were thawed on ice and incubated for 20 minutes with DNA. A heat shock at 42ºC for 1 min was followed by incubation on ice for another 5 min. After addition of 0.9 ml of LB medium, the cells were grown for 1 h at 37ºC. An aliquot of cell suspension was plated on LB agar plates supplemented with the appropriate antibiotic and incubated at 37ºC overnight until colonies became visible.

**Protein expression and whole cell fluorescence measurements.** Expression of prokaryotic homologs was carried out in 96 well plates (2.2 ml deep, round bottom, Thermo Scientific) closed with gas permeable seals (Thermo Scientific). Precultures and expression cultures were grown in 0.7 ml medium. After transformation 4 colonies were picked for each construct and grown as precultures overnight at 37ºC in LB medium. For expression, the preculture was diluted 100 times with the TB medium (12 g peptone, 24 g tryptone, 4 ml glycerol, 2.31 g $KH_2PO_4$, 12.54 g $K_2HPO_4$ and 100 mg ampicilin per liter). The cultures were grown at 37ºC under vigorous shaking to an $OD_{600}$=0.5, cooled down to 25ºC and induced with the following arabinose concentrations: $10^{-4}$, $10^{-3}$, $10^{-2}$ and $10^{-1}$ % (w/v). After overnight expression, the plates with the cultures were centrifuged and the cell pellets resuspended in 100 µl of PBS buffer. 50 µl of cell suspension was transferred to black 96 well plates (Nunc)) and the whole cell fluorescence was recorded in a Tecan GENios Microplate Reader with the Magelan Software ($\lambda_{absorption}$=485 nm, $\lambda_{emission}$=535 nm). To correct for the different cell densities between cultures, the $OD_{600}$ of 10 times diluted cultures were measured in MaxiSorp 96 well plates (Nunc). Specific fluorescence

was calculated by dividing the measured fluorescence value with the measured $OD_{600}$ value.

**In-gel GFP fluorescence detection and Western blot** (from (Geertsma, Groeneveld et al. 2008)). Whole cell expression samples equivalent to ~ 1 mg of total protein were resuspended in 400 µl ice-cold lysis buffer (50 mM K-phosphate pH 7.2, 1 mM MgSO4, 10 % (w/v) glycerol, 1 mM PMSF and 20 µg/ml DNase). After addition of glass beads (300 mg, 0.1 mm diameter), the samples were broken with a Fast prep device (Bio101) two times at 20 s at force position 6 interrupted by 5 min cooling step on ice between the runs. 20 µg of total protein samples, corresponding to 2 µg of total membrane protein, were analyzed by SDS-PAGE. In-gel fluorescence was visualized with an LAS-3000 imaging system (Fujifilm) and analyzed with the AIDA software (Raytest). Immediately after fluorescence visualization, gels were submitted to semidry electroblotting. The standard transfer buffer contained additional 0.1 % SDS to ensure better transfer of membrane proteins. For immunodetection of $His_{10}$-tagged proteins a mouse monoclonal anti His-tag antibody-HRP conjugate was used according to the manufacturer's instructions (Roche).

**Isolation of membrane vesicles.** Cells were harvested and lysed with an Emulsiflex C3 high pressure homogenizer (Avestin) in 50 mM K-phosphate pH 7.4, 150 mM NaCl, 5 % glycerol with addition of 1 mg/ml lysozyme, 20 µg/ml DNase, 1µg/ml leupeptin, 1 µg/ml pepstatin and 1 mM phenylmethyl sulphonil fluoride (PMSF) with at least two passages at 700 bar. The lysate was cleared by low-spin centrifugation (30 min, 11,000 x g, +4ºC). The membranes were pelleted from the supernatant by high-spin centrifugation (1.5 h, 100,000 x g, +4ºC), flash frozen in liquid nitrogen and stored at -80ºC.

**Extraction of membrane proteins from the membrane.** For the small scale extraction tests, membrane proteins were extracted from the cell lysate. The samples were prepared as described for in-gel fluorescence experiments. After lysis, aliquots of 100 µl were solubilized for 1 h at 4ºC in the presence of 1 % DDM (w/v). For analysis, a small aliquot (20 µl) was taken before ultracentrifugation, the remaining sample was centrifuged at 335,000 x g for 15 min at 4ºC in a TLA 100.1 rotor. The pellet after centrifugation was resuspended in 5 M urea. Three samples: total protein, pellet and supernatant after centrifugation were separated by SDS-PAGE and analyzed by in-gel fluorescence and Western blot.

Extraction of membrane proteins from previously isolated membrane vesicles was carried out by mixing 10 ml of solubilization buffer (50 mM K-phosphate pH 7.4, 150 mM

NaCl, 5 % glycerol, 1 % DDM (w/v)) with membranes (10 ml buffer per 1g of membranes). The suspension was incubated for 1.5 h with gentle stirring at 4ºC. To remove unsolubilized membrane fractions, the sample was centrifuged at 100,000 x g for 30 min at + 4ºC in Ti70 or Ti80 rotor (Beckman).

**Purification of prokaryotic ClC proteins.** The supernatant after ultracentrifugation was incubated in batch with equilibrated Ni-NTA agarose (Qiagen) for 1 h. The resin was transferred to a column and washed with 50 mM K-phosphate pH 7.4, 150 mM NaCl, 5 % glycerol, 0.04 % DDM (buffer A) until $OD_{280}$ baseline was reached. To remove nonspecifically bound proteins the resin was washed with buffer A supplemented with 50 mM imidazole. The protein was eluted with buffer A containing 300 mM imidazole. For cleaving of the His-tag HRV 3C protease (GE Healthcare) was added at the ratio 1 mg of protease for 10 mg of protein. After 1-2 h incubation the protein solution was concentrated to less than 500 µl with an Amicon Ultra Centrifugal filter device (MWCO = 50 kDa, Milipore). The concentrated protein was injected onto a gel filtration column (Superdex 200 10/300 GL, Äkta Prime system, GE Healthcare) equilibrated with running buffer (10 mM K-phosphate, 150 mM NaCl, 0.04 % DDM). Size exclusion profile was recorded as absorption at 280 nm. All steps were carried out at 4ºC.

## 2.2 Rm1ClC - a ClC homolog from *Ralstonia metallidurans*

### 2.2.1 Biochemical characterization of Rm1ClC full length protein

In the expression and biochemical screening of prokaryotic ClC proteins, the homolog from the bacterium *Ralstonia metallidurans*, Rm1ClC, showed the most promising behavior. This protein was expressed at high level ~ 2 mg of protein per liter of culture, which is a satisfying amount for getting enough protein for crystallization trials. Besides the high expression it was nearly completely solubilized upon addition of DDM and DM. Size exclusion profiles in DDM and DM confirmed its monodispersity and long-term stability in detergent solution.

All tests for precrystallization screening were performed on a small scale with membrane protein–GFP fusion, where GFP fluorescence had many advantages for the detection of the protein. However, for large scale expression, purification and crystallization the presence of C-terminal GFP is not necessarily beneficial, since the GFP part needs to be cleaved from the membrane protein by limited proteolysis and one more IMAC purification step is required to remove the GFP-His tag part. Therefore, Rm1ClC was cloned into the pBXC3H vector resulting in a C-terminal fusion with a 3C cleavage site and a $His_{10}$ tag (Figure 2.1.3). The expression level and biochemical behavior of the Rm1ClC-His construct were similar to the Rm1ClC-GFP fusion.

#### 2.2.1.1 Large scale expression and purification of Rm1ClC for crystallization

In order to obtain enough protein for crystallization trials, Rm1ClC was expressed and purified on a large scale (10 l culture in TB medium). Protein expression was induced with $10^{-3}$ % arabinose at 25°C overnight. Protein purification was performed as described in the Material and methods chapter. Briefly, Rm1ClC was extracted from the membrane vesicles by DDM or DM and then purified by IMAC. The initial purifications were carried out by using Ni-NTA agarose (Qiagen) as the affinity matrix. Subsequent use of TALON affinity resin (Clontech) proved to be advantageous, because the eluted protein was of higher purity, due to the less non-specific binding to the $Co^{2+}$ chelating resin.

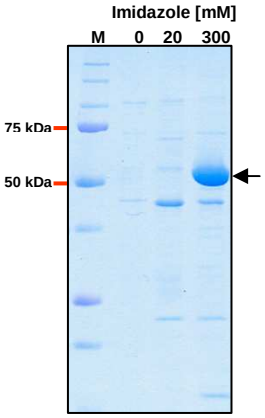



**Figure 2.2.1. SDS-PAGE analysis of Rm1ClC IMAC purification.**
After batch binding to TALON affinity resin, a wash with low imidazole concentration was performed followed by elution of the protein with 300 mM imidazole. The non-specifically bound proteins elute with 20 mM imidazole, while Rm1ClC is still bound to the resin. In the elution fraction Rm1ClC is the major protein band (black arrow), but the minor impurities are still present.

After elution from the affinity resin, the $His_{10}$-tag was removed by digestion with 3C protease. The protein was concentrated and subjected to size exclusion chromatography on a S200 column. Different maltosides were used to assess stability of Rm1ClC and subsequently perform crystallization trials in these detergents. Rm1ClC proved to be stable in all four tested detergents: β-D-dodecyl-maltoside (DDM), β-D-decyl-maltoside (DM), β-D-nonyl-maltoside (NM) and β-D-octyl-maltoside (OM). The protein was eluting at a volume corresponding to a dimer, it was monodisperse and stable after 24 h as confirmed by gel filtration (data not shown). Figure 2.2.2 displays the elution profiles of Rm1ClC in DDM, DM and OM. The corresponding SDS-PAGE analysis of eluted fractions confirms that the highest amount of the protein was eluted in the peak fractions.

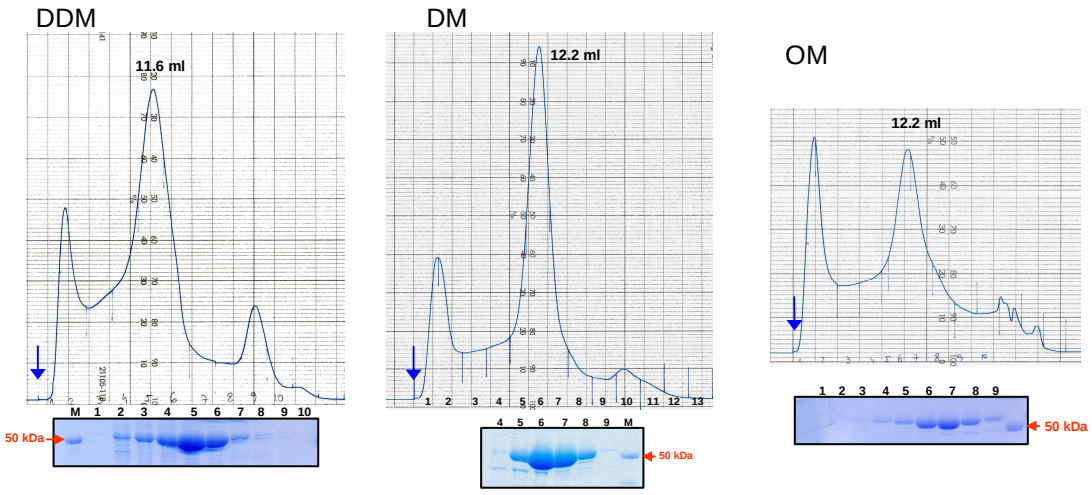



**Figure 2.2.2. Rm1ClC purification in different maltoside detergents.**
Rm1ClC was purified by IMAC and after His-tag removal with the 3C protease additionally purified by size exclusion chromatography on a S200 column. Chromatograms of purifications in DDM, DM and OM are presented together with the corresponding analysis of the fractions by SDS-PAGE. The elution volume in all three detergents suggests a dimer of Rm1ClC protein. The blue arrow indicates the void volume (7 ml).

### 2.2.1.2 Crystallization of Rm1ClC

Rm1ClC was broadly screened for crystallization in the four maltoside detergents DDM, DM, NM and OM and at two different temperatures (4°C and 20°C). Unfortunately, the crystallization screening did not provide any promising results. Precipitation of the protein in the crystallization drops was not observed even at the high precipitant concentrations, which might suggest that the protein is stable under many investigated conditions, but fails to crystallize.

## 2.2.2 Expression and crystallization of the Rm1ClC transmembrane domain

While Rm1ClC showed a very promising biochemical behavior, it failed to crystallize. Although the ultimate goal of the project was the structure determination of a full length ClC protein containing the transmembrane and cytoplasmic domains, I continued with the expression of the two isolated domains independently to overcome the crystallization bottleneck. This strategy would allow to determine the structures of the isolated components in case a flexible linker connecting the two domains prevented crystallization of the full length Rm1ClC. The structural information on the isolated domains would still be very useful for the design of biochemical and functional experiments.

The transmembrane domain of Rm1ClC (Rm1ClC_TM) contains some very interesting features (Figure 2.2.3A): The sequence identity between EcClC and Rm1ClC_TM is 24 % with a sequence similarity of 49 %. The residues building the ion selectivity filter are conserved, except for the serine residue (S107 in EcClC) that coordinates a chloride ion in the central binding site. The N-terminus of Rm1ClC is 19 residues shorter than that of EcClC, leading to a severe truncation or even absence of the A helix preceding the first transmembrane helix. Some loops connecting transmembrane helices are shorter than those found in EcClC, like the ones between helices D and E, and N and O. The loop between helices F and G is, in contrast, three residues longer. Overall, the transmembrane domain of Rm1ClC appears to be very compact and consequently promising for crystallization.

Structural information on EcClC and several ClC cytoplasmic domains enabled the straightforward design of transmembrane domain constructs. In the designed constructs the length of the N- and C-terminus was varied. With respect to the C-terminus, five constructs were generated ending at different positions in the putative R helix and in the linker region between the transmembrane and the cytoplasmic domain (Figure 2.2.3C). The N-terminus of Rm1ClC was truncated by 4, 8 and 12 residues, with the shortest construct starting at the first transmembrane helix B (Figure 2.2.3B).

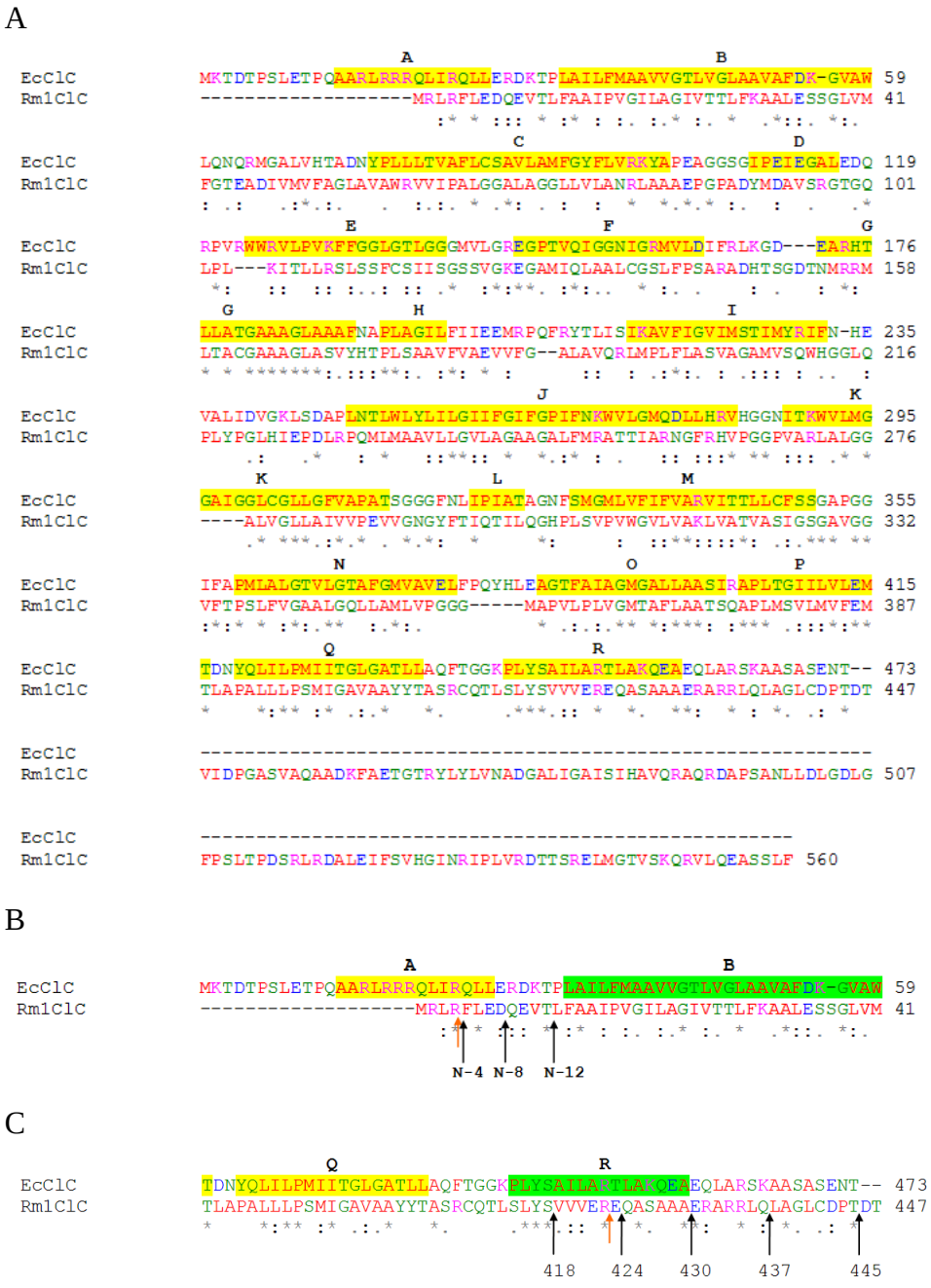



**Figure 2.2.3. Sequence alignment between EcClC and Rm1ClC and design of Rm1ClC transmembrane domain constructs.**
**(A)** Sequence alignment between EcClC and Rm1ClC generated in ClustalW. The residues building the helices in EcClC are marked yellow and the helices are labeled from A to R. The identical residues are labeled with "*", the similar ones with ":". The N-terminus of Rm1ClC is significantly shorter than the one in EcClC. The long C-terminal part in Rm1ClC is a cytoplasmic CBS domain. **(B)** The truncations of N- and C-termini of Rm1ClC were based on the sequence alignment and the known structure of EcClC. At the N-terminus, 3 truncations lead to 4, 8 and 12 aa residues shorter proteins. The N-12 construct starts at the first putative transmembrane helix (marked green). **(C)** At the C-terminus, five truncations were constructed. Two of them (418 and 424) terminate within the R helix (marked green). Construct 430 ends at the last amino acid belonging to the R helix of EcClC. Construct 437 ends at the position equivalent to the EcClC construct used for crystallization. Construct 445 ends at the position corresponding to the native C-terminus of EcClC. Red arrows indicate the sites of tryptic digestion experimentally determined by a combination of limited proteolysis on Rm1ClC and mass spectrometry.

The goal of the design of constructs with different length was to identify the ones that give the highest crystal and diffraction quality. Since long and floppy termini can inhibit crystallization or influence crystal packing, it is generally helpful to find a minimal stable construct for crystallization that leads to improved diffraction.

### 2.2.2.1 Cloning and expression of Rm1ClC transmembrane domain constructs

For cloning of Rm1ClC transmembrane domain constructs FX cloning was applied. DNA fragments were cloned into the pBXC3GH vector to generate Rm1ClC_TM–GFP fusions. Test expression and quality control were investigated as described for previously cloned ClC homologs.

All truncated constructs were expressed at a similar or even higher level compared to the full length protein. Interestingly, when trying to express the transmembrane domain only as a $His_{10}$ tagged protein (pBXC3H vector), the expression was almost undetectable by Western blot. For this reason, Rm1ClC_TM-GFP fusions were subsequently used for large scale expression and purification.

The quality control assay of overexpressed transmembrane domain constructs showed no presence of aggregated protein at all investigated inducer concentrations. All constructs were completely extractable with decyl-maltoside and monodisperse in size exclusion chromatography as illustrated in the SEC profile of the Rm1ClC transmembrane domain construct 445 (Rm1ClC_TM445) in Figure 2.2.4.

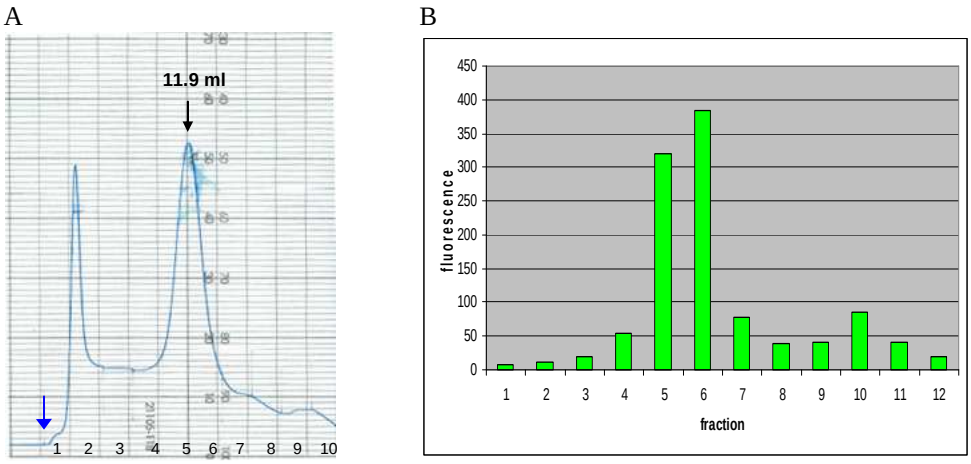



**Figure 2.2.4. Size exclusion chromatography of Rm1ClC_TM445-GFP fusion.**
Purified GFP fusion of Rm1ClC_TM445 was analyzed by SEC on a S200 column. The chromatogram (A) shows the absorbance at 280 nm during the run, the histogram (B) presents fluorescence values in each eluted fraction. The blue arrow is indicating the void volume (7 ml). The chromatogram shows a symmetric peak at 11.9 ml, the volume corresponds to a well folded protein at the approximate size of a dimer. The fluorescence measurement confirms the monodispersity of Rm1ClC_TM445-GFP fusion.

### 2.2.2.2 Large scale expression and purification of Rm1ClC_TM constructs

Large scale expression of the Rm1ClC transmembrane domain – GFP fusion was performed in the same way as described for the expression of the full-length protein. Bacteria were grown in 10 l cultures in TB medium and protein expression was induced with $10^{-3}$ % arabinose at 25°C overnight.

Protein purification was performed as described in the Material and methods chapter. Briefly, Rm1ClC_TM-GFP fusions were extracted from isolated membrane vesicles by DM and then purified by IMAC (Figure 2.2.5). TALON affinity resin (Clontech) was used in all experiments since it has been previously observed to be more efficient in yielding pure protein after elution.

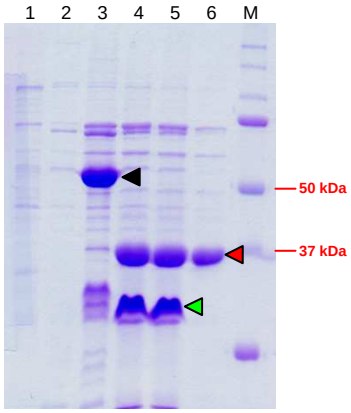



**Figure 2.2.5. SDS-PAGE analysis of the Rm1ClC transmembrane domain purification.**
The Rm1ClC_TM430-GFP fusion was extracted in DM and purified with TALON affinity resin. After batch binding, the resin was washed with low imidazole concentrations to remove non-specifically bound proteins (sample 1 – 0 mM, sample 2 – 20 mM imidazole) and the fusion protein was eluted with 300 mM imidazole (sample 3). To remove the GFP-$His_{10}$ part, the protein was cleaved with the 3C protease for 1 h (samples 4 and 5). The GFP-$His_{10}$ was subsequently separated from the Rm1ClC transmembrane domain by binding to Ni-NTA agarose (sample 6). Black triangle – Rm1ClC_TM430-GFP fusion; Red triangle – Rm1ClC_TM430; Green triangle – GFP-$His_{10}$.

To cleave off the GFP-$His_{10}$ part from the Rm1ClC transmembrane domain, the protein was digested with 3C protease followed by an additional IMAC purification in order to remove GFP-$His_{10}$ and to obtain pure Rm1ClC transmembrane domain. Apart from GFP-$His_{10}$, this step also removes some impurities, which were co-purified on the first metal affinity column.

Size exclusion chromatography was used as an additional purification step and to investigate the homogeneity and stability of Rm1ClC_TM in different maltoside detergents. The Rm1ClC transmembrane domain was stable in all tested detergents: β-D-decyl-maltoside (DM), β-D-nonyl-maltoside (NM) and β-D-octyl-maltoside (OM). The protein was eluting at the expected elution volume for a dimer, was monodisperse and stable after 24 h reinjection (data not shown). Figure 2.2.6 illustrates the elution profiles of Rm1ClC_TM430 in DM, NM and OM. The SDS-PAGE analysis after purification in DM confirms that the highest amount of the protein was eluted in the peak fractions. The eluted protein was also pure enough for crystallization.

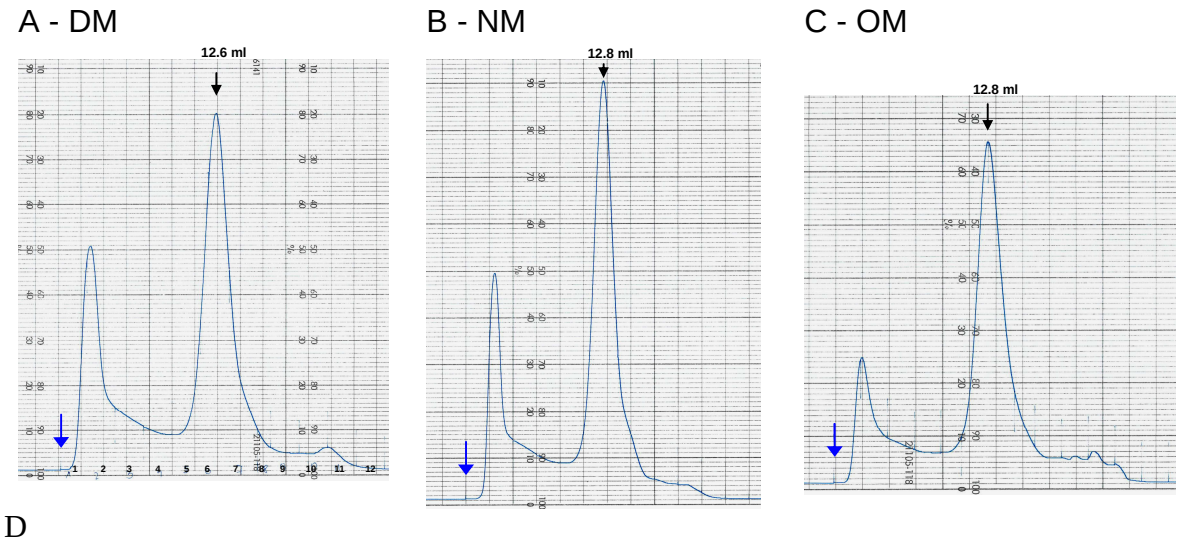



**Figure 2.2.4 Size exclusion chromatography of Rm1ClC transmembrane domain in different maltoside detergents.**
After IMAC purification in DM and removal of GFP-$His_{10}$, the Rm1ClC_TM430 was purified by size exclusion chromatography on a S200 column equilibrated with different detergents (A – decyl-maltoside, DM; B – nonyl-maltoside, NM; C – octyl-maltoside, OM). The exclusion profile in all three detergents indicates a monodisperse protein with the expected elution volume for a dimer (12.6 ml). The blue arrow indicates the void volume (7 ml). The SDS-PAGE analysis of the fractions from SEC in DM (D) shows that the Rm1ClC transmembrane domain is mostly found as pure protein in the main elution peak. Fractions 6, 7 and 8 were pooled and subsequently used for crystallization.

### 2.2.2.3 Crystallization of Rm1ClC_TM constructs

Rm1ClC_TM constructs were broadly screened for crystallization in three maltoside detergents (OM, NM and DM) and at two different temperatures (4°C and 20°C). The first trials, which were set up with the protein construct 430 in DM, showed crystals under many conditions at 4°C with PEG400 as precipitant. Refined screening for this and other constructs gave similar results with respect to the crystallization conditions. The change of detergent from DM to OM led to a narrowing of the crystallization conditions with respect to salt type and pH range. At this stage, the diffraction quality of different crystal forms was assayed. Table 2.2.1 summarizes crystallized Rm1ClC_TM constructs, their crystallization condition and X-ray diffraction properties.

The crystallization of six different Rm1ClC_TM constructs allowed to obtain crystals growing under similar conditions, most of which showed low resolution diffraction, extending in the best cases to resolution of 6 to 7 Å.

| **Crystallized Rm1ClC_TM construct** | **Detergent** | **Crystallization condition** | **Diffraction properties** | **Unit cell parameters (space group, unit cell axis in Å, unit cell angles different than 90°)** |
|---|---|---|---|---|
| 418 | NM | 15 % PEG400, 0.5 M KCl, pH 6.5 | Max to 6 Å | $P2_1$<br>82.1 202 87.4 β=104° |
| 424K_3C | DM | 25% PEG400, 0.5M KCl, pH 5.5 | Anisotropic diffraction<br>6 Å good direction<br>7.5 Å bad direction | $P2_1$<br>87 208 87 β =103° |
| | NM | 25% PEG400, 0.5M KCl, pH 8.4 | Max to 7 Å | |
| | OM | 25% PEG400, 0.5M KCl, pH 8.4 | Max to 7 Å | |
| 424K_LysC | DM | 25% PEG400, 0.5M KCl, pH 5.5 | Anisotropic diffraction<br>4.5 Å good direction<br>6.5 Å bad direction | $P2_1$<br>84.94 207.1 87.25 β =103.2° |
| | NM | 25% PEG400, 0.5M KCl, pH 8.4 | Max to 9 Å | |
| | OM | 25% PEG400, 0.5M KCl (KBr or NaBr)<br>pH 8.4 | Max to 9 Å | |
| 430 | DM | 15-25% PEG 400, various salts<br>broad pH range | Max to 6.5 Å | P2<br>92.3 143.7 122.1 β = 99.8° |
| | NM | 24 % PEG400, 0.1 M Na,K-tartrate, pH 5.5<br>20 % PEG400, 1 M NaCl, pH 8.4 | Max to 7.5 Å | |
| | OM | 25% PEG 400, 0.5 M KCl, pH 9.4 | Max to 7.5 Å | |
| 437 | DM | 15-25% PEG 400, 0.5 M KCl<br>pH 5.5, 6.5, 7.4 | Max to 7.6 Å | P2<br>119.3 110.8 125.2 β=104° |
| N-8_424K_LyS | DM | 25 % PEG400, various salts<br>pH 7.4, 8.4, 9.4 | Max to 8 Å | |

**Table 2.2.1. Summary of crystallized Rm1ClC transmembrane domain constructs.**
Six different RM1ClC_TM constructs were crystallized in three different detergents at 4ºC. The crystallization conditions for the best crystals for each construct are given. The precipitant in all crystallization condition was PEG400. The pH for crystallization varied from 4.5 to 9.5. Different salts in the reservoir solution contributed to crystal formation, but 0.5 M KCl was the most preferable one. All crystals diffracted to low resolution, in most of the cases to around 7 Å. The crystals of the construct 424K_LysC had the most promising diffraction properties. Although the diffraction was anisotropic, it reached in one direction resolution of 4.5 Å. For some of the crystals the full data sets were collected. The data was processed; the unit cell parameters are presented in the table. All analyzed crystals belong to a primitive monoclinic space group, with some difference in unit cell constants.

For some crystals data collection and crystallographic characterization were performed. All analyzed crystals were of primitive monoclinic space group (P2 or $P2_1$), but had different unit cell parameters and Matthews coefficients depending on the construct.

The length of the N- and C-termini in different constructs did not play a very significant role for crystal quality (except in one case, see later), as for all tested constructs crystals diffracting to 7 Å were identified under certain crystallization conditions (Figure 2.2.8A). Surprisingly, strong dependence of the diffraction quality on the detergent used for crystallization was also not observed. In crystallization of membrane proteins, the length of the hydrophobic chain of the detergent often plays a very important role for crystal quality. The optimal chain length depends on the membrane protein structure and has to be determined experimentally. Detergents with short acyl-chain are often favorable, since they cover less protein surface and therefore leave more places for crystal contacts. Many structures of membrane proteins were successfully determined by using octyl-maltoside and octyl-glucoside for crystallization. Unfortunately, crystallization of Rm1ClC_TM in octyl-maltoside did not improve diffraction properties compared to its crystallization in decyl-maltoside. The transmembrane domain of Rm1ClC was previously described as being compact and thus promising for crystallization. However, this very compact architecture could also be disadvantageous particularly when hydrophilic loops between the transmembrane helices are too short to make well defined crystal contacts, which might be the case in Rm1ClC_TM.

The best diffraction was obtained with the construct 424K_LysC crystallized in 5 mM DM, 25 % PEG400, 0.5 M KCl, pH 5.5. Other transmembrane domain constructs gave crystals under this or similar conditions of lower quality. Therefore, it was assumed that the length of the C-terminus in the 424K_LysC construct plays a central role for crystal packing. In this construct a lysine residue was introduced after the amino acid 424 and before the recognition site for the 3C protease (Figure 2.2.7). In that way, it was possible to remove the residues of the 3C recognition site by digestion with the protease LysC, which cleaves after lysine residue. The complete digestion was also detectable on SDS-PAGE and later confirmed by mass spectrometry.

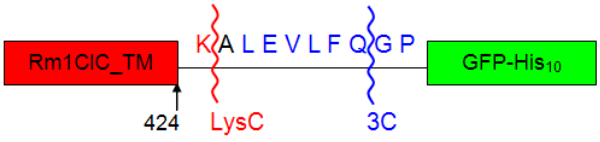



**Figure 2.2.7. Schematic representation of Rm1ClC transmembrane domain 424K_LysC construct.** After the amino acid residue 424 and before the recognition site for the 3C protease (blue) a lysine residue is introduced by cloning (red). Following GFP-$His_{10}$ removal by digestion with the 3C protease, digestion with LysC removed the remaining residues of the 3C recognition site.

The crystals of Rm1ClC_TM construct 424K_LysC diffracted anisotropically to different resolutions in different regions of reciprocal space. In the direction of poor diffraction the crystal diffracted to about 6.5 Å, and in the direction of good diffraction to 4.5 Å resolution (Figure 2.2.8B). Anisotropic diffraction usually occurs if the crystal packing is more ordered in one direction than in others. Frequently, the directional dependence coincides with the reciprocal cell axes a*, b*, and c*.

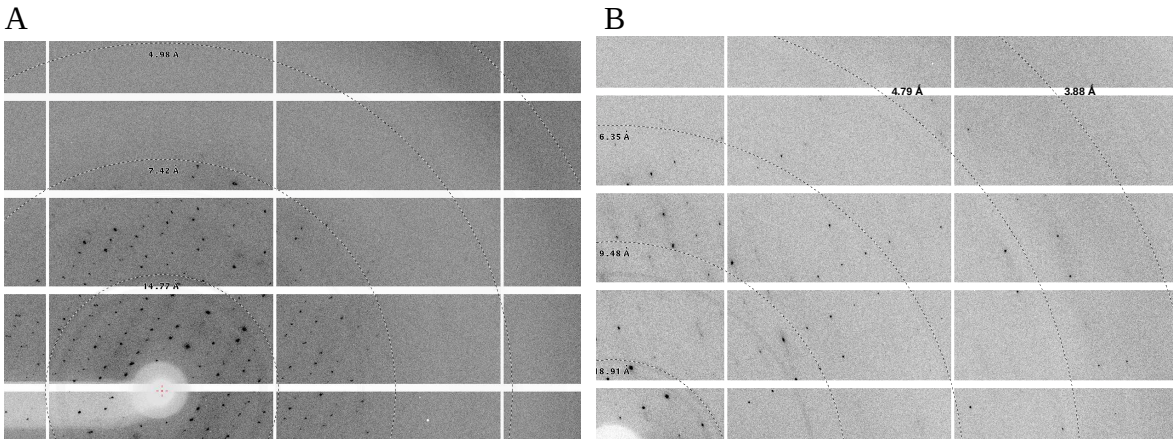



**Figure 2.2.8. Diffraction images of Rm1ClC_TM crystals.**
A) Crystal of the construct 430 grown in 26 % PEG400, 0.15 M NaCl, pH 6.5 and OM detergent. Resolution rings are indicated. The crystal diffracted isotropically to 7.5 Å. Many crystals from other TM constructs and crystallization conditions behaved in a similar way.
B) Crystal of the construct 424K_LysC in 25 % PEG400, 0.5 M KCl, pH 5.5 and DM. The crystal diffracted anisotropically. In one direction the diffraction was observed to ~ 6.5 Å, and in the other direction to ~ 4.5 Å resolution. The diffraction image is from the dataset which was used for molecular replacement.

The 424K_LysC data was processed. The crystal belongs to the space group $P2_1$ with the following cell parameters: a = 84.94 Å, b = 207.1 Å, c = 87.25 Å, β =103.2°. The number of molecules in the asymmetric unit (AU) was estimated by calculating the Matthews coefficient (Matthews 1968), suggesting 4 molecules per AU and a solvent content of 72 % (Table 2.2.2).

| N molecules/AU | Matthews Coeff | % solvent |
|---|---|---|
| 1 | 17.74 | 93.07 |
| 2 | 8.87 | 86.14 |
| 3 | 5.91 | 79.21 |
| 4 | 4.43 | 72.28 |
| 5 | 3.55 | 65.35 |
| 6 | 2.96 | 58.42 |
| 7 | 2.53 | 51.48 |
| 8 | 2.22 | 44.55 |
| 9 | 1.97 | 37.62 |
| 10 | 1.77 | 30.69 |

**Table 2.2.2. Calculation of Matthews Probabilities and Solvent Content of 424K_LysC crystal.**
The Matthews coefficient and solvent content were calculated from the unit cell parameters and the molecular weight of the molecule in the unit cell (43881 Da) with the CCP4 program MATTEWS_COEF. The number of molecules in the AU with the corresponding Matthews coefficient and solvent content is presented. Four molecules in the AU would give a solvent content of 72%, which is in the expected range for membrane protein crystals.

As expected, the completeness of the data was influenced by anisotropic diffraction being 95 % complete only to a resolution of 6 Å. Table 2.2.3 illustrates data statistics after scaling.

| Resolution limit | Redundancy | Completeness (%) | $R_{merged}$ (%) | I/σ |
|---|---|---|---|---|
| 10.0 | 5.6 | 86.0 | 2.4 | 61.31 |
| 8.0 | 6.6 | 99.4 | 3.2 | 48.64 |
| 7.0 | 7 | 98.6 | 5.5 | 30.31 |
| 6.0 | 6.5 | 95.6 | 15.0 | 12.98 |
| 5.5 | 6.5 | 90.5 | 30.7 | 7.19 |
| 5.0 | 6.1 | 87.4 | 38.5 | 5.91 |
| 4.8 | 5.8 | 81.3 | 45.9 | 5.07 |
| 4.6 | 5.5 | 79.3 | 57.0 | 4.20 |
| 4.4 | 5.4 | 75.0 | 65.7 | 3.81 |
| 4.2 | 5 | 69.9 | 112.9 | 2.25 |
| total | 5.9 | 85.5 | 7.1 | 16.06 |

**Table 2.2.3. Statistics for the data from 424K_LysC crystal after scaling in XSCALE.**

### 2.2.2.4 Molecular replacement

The dataset obtained from the crystal of the 424K_LysC construct was of sufficient quality for molecular replacement with the program Phaser. As search model the EcClC structure, encompassing residues 20-451, was used. Whereas an initial trial with the EcClC dimer as search model failed, another search with a single EcClC subunit as search model was successful. When searching with a single subunit, three out of four predicted copies in the asymmetric unit were identified. Two of them had a similar dimeric arrangement as observed in the EcClC structure. A careful analysis by superposition showed that the relationship between the monomers in the molecular replacement solution does not obey a perfect two fold symmetry as in EcClC but it is not clear whether this is due to the limited resolution of the data (Figure 2.2.10A). Potentially this is the reason why the molecular replacement with EcClC dimer as search model was unsuccessful. In a subsequent search the dimer identified in the first run was used as search model. After increasing the tolerance for clashes in the packing from 10 to 25, the second dimer was found. Two dimers interact via the intracellular side of the protein where also the overlaps between the two models were observed (Figure 2.2.10B). After removal of four residues from the N-terminus and two residues in the loop between the H and I helices from the search model, a final search with the truncated model gave a strong solution with very good parameters (translation function Z-score (TFZ) of 18 and the refined log-likelihood gain (LLG) of 394) indicating a correct packing of the model The validity of the molecular replacement

solution was further examined by comparing self rotation functions calculated from the experimental data with the calculated rotation function from the obtained solution (Figure 2.2.9). The projections of κ=180º sections look very similar indicating that the obtained molecular replacement solution is most probably correct.

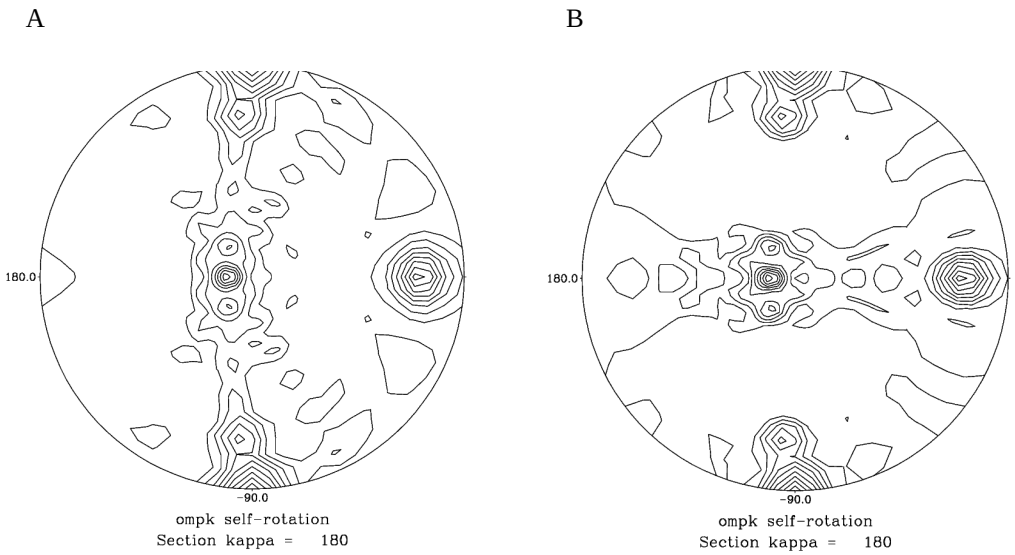



**Figure 2.2.9. Self rotation functions for the experimental data of 424K_LysC crystal (A) and the molecular replacement solution (B).** Shown is the kappa section at 180º.

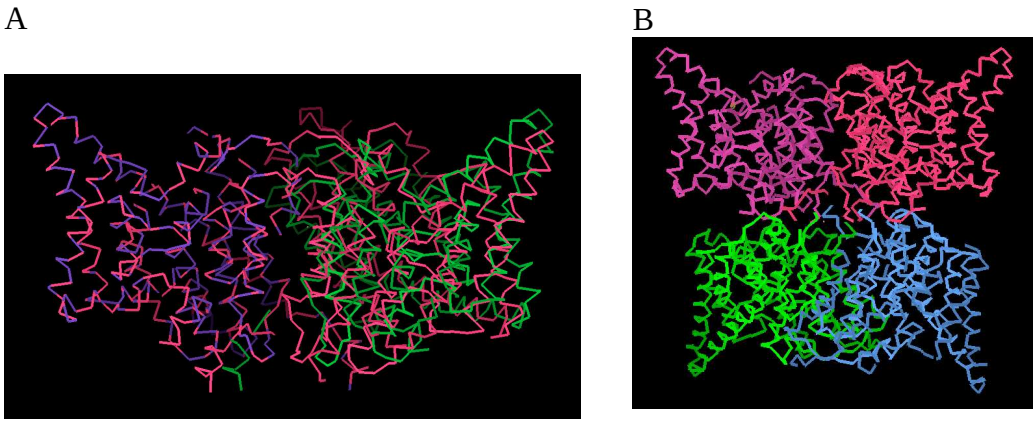



**Figure 2.2.10. Molecular replacement solution from 424K_LysC construct.**

**(A)** Superposition of EcClC dimer with the dimer found in the crystal of 424K_LysC construct by molecular replacement. The chain A of one dimer is superimposed on chain B of the other dimer (Secondary-structure matching (SSM) in Coot). The subunits in the dimer obtained by molecular replacement (violet and green) are not related by a perfect 2 fold. **(B)** Arrangement of four molecules in the asymmetric unit of 424K_LysC crystal. The interaction between two dimers is formed with the intracellular loops.

The crystal packing of the molecular replacement solution revealed some interesting features (Figure 2.2.11). The crystal contacts in the direction of the three cell axes were not of the same extent. The strongest contacts are observed parallel to the x axis, where the interaction between the dimers in the asymmetric unit is also formed. In orientations parallel to y and z axes no direct contacts between the molecules are observed suggesting detergent mediated crystal contacts. The unequal crystal packing interactions might be the reason for the anisotropic diffraction of the crystal. Based on this observation it would be expected that the diffraction quality could be improved by crystallization in detergents with shorter aliphatic chains or by applying post-crystallization treatments like for example dehydration.

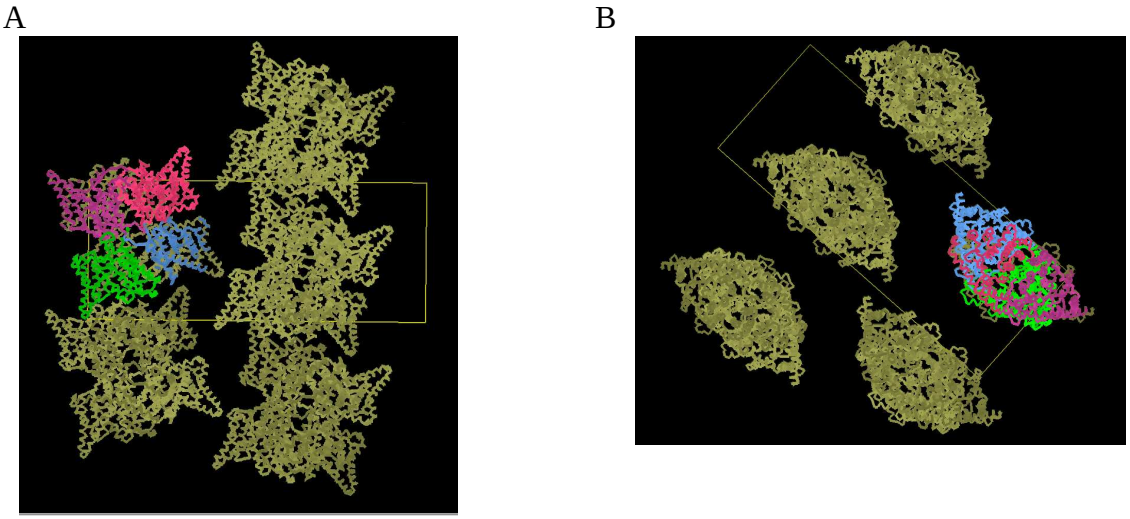



**Figure 2.2.11. Crystal packing in the molecular replacement solution of the 424K_LysC construct.**
**(A)** A view along x axis. The crystal contacts between the asymmetric units are made by the extracellular loops
**(B)** A view along z axis. The molecules are separated by large gaps, suggesting detergent mediated crystal contacts.

### 2.2.2.5 Summary

Although the biochemical properties of the Rm1ClC full length protein make the protein a promising candidate for structure determination, no crystallization was identified, despite the broad screening of conditions. In order to improve the chances for crystallization, the transmembrane domain of Rm1ClC was expressed independently. While the isolated transmembrane domain showed similar biochemical behavior and stability as the full length protein, it readily crystallized. Unfortunately, the crystals were not of sufficient quality for structure determination of Rm1ClC, a feature that is frequently observed for membrane proteins. To overcome this bottleneck in membrane protein crystallography the

production of monoclonal antibodies against Rm1ClC was launched and specific binders for future co-crystallization experiments were selected (Appendix 4.1).

### 2.2.2.6 Material and methods

**Large scale expression and purification of the Rm1ClC full length protein and the transmembrane domain.** For large scale expression 2 l baffled shake flasks were filled with 600 ml TB medium supplemented with 100 µg/ml ampicilin. Usually 9.6 l culture was used for production of the protein for crystallization, resulting in a yield of about 3 mg of purified protein. Cells were grown at 37ºC until an $OD_{600}$ of 0.6 and subsequently cooled to 25ºC. Protein expression was induced with 0.001 % arabinose for 14 - 15 hours at 25ºC.

Cells were harvested by centrifugation at 4000 rpm (rotor Sorvall H6000A) for 20 min at 4ºC. The cell pellet was resuspended in 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 5 % glycerol (buffer B) with the addition of 1 mg/ml lysozyme, 20 µg/ml DNase, 1µg/ml leupeptin, 1 µg/ml pepstatin and 1 mM PMSF. The cells were lysed with an Emulsiflex C3 high pressure homogenizer (Avestin) with at least two passages at 700 bar. The lysate was cleared by low-spin centrifugation at 11,500 × g for 40 min at 4ºC. Membranes were isolated by ultracentrifugation at 45,000 rpm for 1.5 h at 4ºC (Ti70 rotor, Beckman). For storage, membranes were homogenized in buffer B in 1/10 of the extraction volume, flash frozen in liquid nitrogen and stored at -80ºC.

All purification steps were carried out at 4ºC. Membranes were resuspended in buffer B (10 ml buffer/1 g membrane) and solubilized by addition of detergent (1 % for DDM, 1.25 % for DM) for 1.5 h. The unsolubilized membrane fraction was removed by centrifugation at 45,000 rpm for 30 min at 4ºC (Ti70 rotor, Beckman). The supernatant was supplemented with 3 mM imidazole to reduce nonspecific binding to the metal affinity resin.

The protein was purified by IMAC using TALON affinity resin (Clontech). All buffers used for purification contained detergent - 0.8 mM DDM or 5 mM DM. 10 ml of equilibrated resin was used for batch binding for 1h. Afterwards, the resin was transferred to a column and washed with the buffer B until baseline of $OD_{280}$ was reached. To remove nonspecifically bound proteins the resin was washed with the buffer B supplemented with 20 mM imidazole. The protein was eluted with the buffer B containing 300 mM imidazole.

To cleave off the $His_{10}$-tag or remove the GFP-$His_{10}$ part, the protein was digested for 2 h with HRV 3C protease (GE Healthcare) at a ratio of 1 mg of protease for 10 mg of fusion protein. During the digestion of Rm1ClC-TM-GFP fusions, the solution was dialyzed against buffer B containing detergent at a concentration of 1.2 × CMC to remove

imidazole. After digestion of Rm1ClC-TM-GFP fusions, the solution was incubated for 15 min with 2 ml Ni-NTA agarose (Qiagen) to remove the GFP-$His_{10}$ and the His-tagged 3C protease.

The cleaved protein was concentrated to a volume of ~ 400 µl using an Amicon Ultra Centrifugal filter device (MWCO = 50 kDa, Milipore). The concentrated sample was subjected to size exclusion chromatography on a Superdex S200 column (GE Healthcare) equilibrated in buffer containing 10 mM Tris-HCl pH 7.4, 150 mM NaCl and a detergent at the following concentration: 0.8 mM DDM, 5 mM DM, 20 mM NM or 30 mM OM. In that way it was possible to exchange the detergent used for extraction and IMAC with the detergent needed for crystallization.

**Crystallization of Rm1ClC full length protein and transmembrane domain.** The purified protein was concentrated to 10 mg/ml and used for crystallization. Crystallization screening was performed at the NCCR High-Throughput Nanoliter Crystallization Facility. All crystals were obtained by sitting drop vapor diffusion at 4ºC by mixing protein with the reservoir solution in 1:1 ratio. After screening, the crystallization conditions were scaled up to 2 µl drops and refined by hand (Cryschem-plates, Hampton Research). The crystals of Rm1ClC transmembrane domain were grown under conditions that contained 20 - 30 % PEG400 as precipitant, 50 mM to 1 M salt (e.g. NaCl, KCl, $CaCl_2$, $MgCl_2$, $NH_4Cl$, K, Na-tartrate etc.) and a pH varying from 4.5 to 9.4. In some cases, the protein was supplemented with 0.5 mg/ml of *E. coli* polar lipids (Avanti Polar Lipids, Inc), which did not improve the crystal quality.

For cryoprotection, crystals were transferred into mother liquor containing an increasing amount of PEG400 to a final concentration of 35 %, flash frozen in liquid propane and stored in liquid nitrogen.

**Crystallography.** Data was collected on frozen crystals on the X06SA beamline at the Swiss Light Source of the Paul Scherer Institute on a PILATUS detector (Dectris). The data were indexed, integrated and scaled with XDS (Kabsch 1993) and further analyzed and processed with CCP4 programs (Bailey 1994).

The dataset of the transmembrane domain crystal 424K_LysC was used for molecular replacement with the program PHASER (McCoy, Grosse-Kunstleve et al. 2007). As a search model the structure of a single EcClC subunit (truncated model 20-451 aa) was used. PHASER was run in the automated molecular replacement mode. The resolution was limited to 6 Å. In order to identify the right solution the default packing criterion was changed from 10 to 25. The coordinate files and electron density maps were visualized in the graphical program Coot (Emsley and Cowtan 2004). The self rotation function was calculated using the program polarrfn of the CCP4 package.

### 2.2.3 Crystallization and structure determination of the Rm1ClC cytoplasmic domain

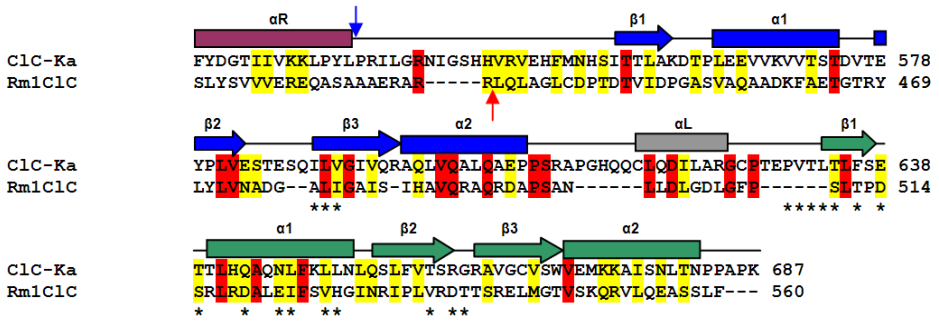
The main criterion for the selection of prokaryotic ClC homologs was the presence of cytoplasmic CBS domains. These prokaryotic proteins could thus serve as valuable models for providing structural and functional information on regulatory mechanisms conferred by these domains. Cytoplasmic domains of prokaryotic ClCs are usually of smaller size than their eukaryotic counterparts. The size difference is due to a shorter linker region between two CBS subdomains and a shorter C-peptide. In the selected set of ClC homologs the average length of the CBS domain, starting from the first amino acid after the R-helix was 147 residues. The degree of conservation when compared to the eukaryotic domains, but also within the prokaryotic domains, is very low. Therefore, it is difficult to predict the role of CBS domains in prokaryotic ClC proteins based on the existing structural and functional information on eukaryotic cytoplasmic domains. To reveal whether prokaryotic cytoplasmic domains share some of the eukaryotic structural and regulatory features or whether they have some unique characteristics they need to be structurally and functionally investigated.

The cytoplasmic domains of ClC proteins can usually be expressed in isolation as soluble proteins, which facilitates their crystallization and structure determination. In addition, biochemical assays for identification of regulating ligands are performed easer on isolated cytoplasmic domains than on labile, detergent solubilized membrane proteins. The obtained results on soluble domains can be valuable for later purification, crystallization and biochemical characterization of the full length protein.

To gain an insight into the structure of the cytoplasmic domain of the ClC homolog from *Ralstonia metallidurans*, this protein component was cloned and expressed independently from its transmembrane part. The expression construct included residues 436–560. Like the full length Rm1ClC, the cytoplasmic domain is very compact and with only 125 amino acids the smallest ClC CBS domain studied so far.

Figure 2.3.1 shows sequence alignments between the cytoplasmic domain of Rm1ClC with the domain of the human channel ClC-Ka (A) and the domain of the prokaryotic homolog PtClC (B), the structures of which have been previously determined. The level of sequence conservation is low. The Rm1ClC and ClC-Ka CBS domains share 15 % of identical and 30 % of similar residues; with respect to the PtClC domain the sequence identity is 13 % and the similarity 25 %. The conserved residues are distributed over the whole domain. Interestingly, the homology between domains of the two prokaryotic ClC proteins is not higher than between the prokaryotic and the eukaryotic ClC protein. The CBS1 subdomain of Rm1ClC is slightly more similar to ClC-Ka and CBS2 to PtClC.

A




B

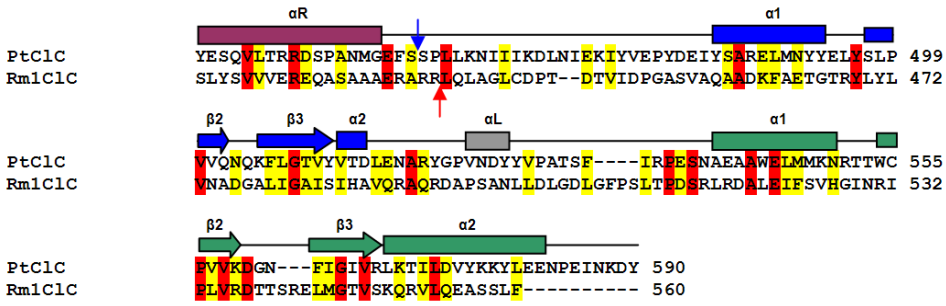



**Figure 2.3.1. Structure-based sequence alignment between Rm1ClC cytoplasmic domain and the cytoplasmic domain of ClC-Ka (A) and PtClC (B).**
Identical residues are highlighted in red, similar in yellow. Secondary structure of ClC-Ka and PtClC domain are indicated above the sequences. The R helix preceding the cytoplasmic domain is included in the alignment and colored in violet. The secondary structure elements of CBS1 subdomain are colored in blue and of CBS2 in green. The helix in the linker region is colored in grey. The first residues of the crystallized ClC-Ka and PtClC products are designated with a blue arrow, and for the Rm1ClC domain with a red arrow. The residues involved in the dimer interface of ClC-Ka domain are labeled with a star.

### 2.2.3.1 Expression and purification of the Rm1ClC cytoplasmic domain

The Rm1ClC cytoplasmic domain was overexpressed as a C-terminal $His_{10}$-tag fusion under the control of the arabinose promoter. The standard purification protocol was applied including IMAC as the first purification step, followed by His tag removal with protease digestion and SEC as the final purification step. Expression and purification resulted in a yield of about 10 mg of purified Rm1ClC domain per liter of expression culture.

From the elution volume on Superdex 200 column it was difficult to estimate the oligomeric state of Rm1ClC CBS domain in solution, because both a monomer and a dimer would be consistent with the observed elution volume.

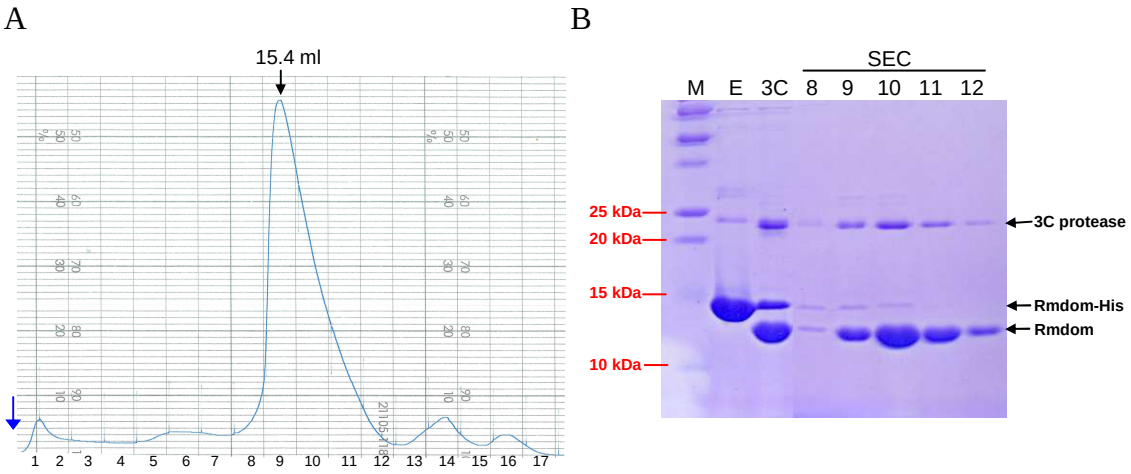



**Figure 2.3.2. Purification of the Rm1ClC cytoplasmic domain.**
After IMAC purification and removal of the His tag with 3C protease, the concentrated protein was subjected to SEC on S200 column (A). The protein eluted as a single monodisperse peak at 15.4 ml. The peak is not symmetrical and shows tailing due to the large amount of protein loaded onto the column. The purification was analyzed by SDS-PAGE (B). Samples: E – protein eluted from the IMAC column, 3C – protein after protease cleavage, 8, 9, 10, 11 and 12 – peak fractions from the size exclusion chromatography. The cleavage of the His tag was not complete. The 3C protease co-migrates with Rm1ClC domain on a S200 column.

### 2.2.3.2 Oligomeric state of the Rm1ClC cytoplasmic domain

To gain an insight into the oligomeric organization of the Rm1ClC cytoplasmic domain its sedimentation behavior was investigated by analytical ultracentrifugation (AUC).

This method allows to determine the properties of the protein in solution including molar mass, gross shape, and the heterogeneity of the sample. Two basic types of experiments are possible: sedimentation velocity and sedimentation equilibrium centrifugation. In sedimentation velocity experiments a high centrifugal force is applied and the time course of the sedimentation process is analyzed, allowing determination of the sedimentation coefficient S and the approximate estimation of the molecular weight (Schuck 2000). In sedimentation equilibrium the application of a low centrifugal force permits the diffusion to balance the sedimentation such that a time-invariant equilibrium gradient can be observed. Because the concentration distribution at equilibrium depends only on molecular mass, the exact molecular mass can be calculated with the known parameter S that was previously determined from sedimentation velocity experiments.

The Rm1ClC domain was analyzed by sedimentation velocity experiments. The experiments were carried out at three different protein concentrations (35, 70 and 150 mM), in TBS buffer at 4°C. The sedimentation profile that looked similar for all protein concentrations is presented in Figure 2.3.3. The measured sedimentation coefficient was 1.1 S. Since the measured S depends on the buffer composition, temperature and protein

concentration, it is necessary to extrapolate it to the value in water at 20°C at infinite dilution - $S^{o}_{20,W}$. The condition independent $S^{o}_{20,W}$ allows indirect estimation of molecular weight. From the calculated $S^{o}_{20,W}$ of 1.77, it can be concluded that the Rm1ClC cytoplasmic domain exists as a monomer in solution at all investigated concentrations.

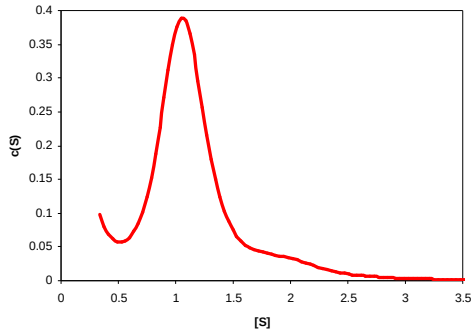



**Figure 2.3.3. Analytical ultracentrifugation of Rm1ClC cytoplasmic domain.** The distribution of sedimentation coefficient as calculated from the sedimentation velocity experiment is plotted against the sedimentation coefficient S. Rm1ClC domain sediments at 1.1 S. The protein concentration in this experiment was 70 µM.

A similar sedimentation behavior was found for the cytoplasmic domain of the prokaryotic ClC homolog PtClC, which was a monomer in solution while all investigated cytoplasmic domains of eukaryotic ClCs showed a dimeric organization in solution. These biophysical properties already hint at potential differences that might exist in the structure and function of prokaryotic and eukaryotic ClC proteins.

### 2.2.3.3 Crystallization of the Rm1ClC cytoplasmic domain

The purified Rm1ClC cytoplasmic domain was broadly screened for crystallization at different protein concentrations (5-10 mg/l) and at two different temperatures (4ºC and 20ºC). The crystals were obtained in collaboration with Sara Savaresi. Interestingly, the protein crystallized at both 4ºC and 20ºC only in two conditions: 35-45 % PEG400, 200 mM $CaCl_2$, pH 7.4 or 9.4. The crystals had a cubic shape with a size of around 20 µm only. At a synchrotron source these small crystals showed very high diffraction to a resolution of 1.6 - 1.7 Å (Figure 2.3.4).

The Rm1ClC domain crystals were of the space group $P2_12_12$ with a small unit cell (a = 40.5 Å, b = 88.2 Å, c = 39.1 Å and α=β=γ= 90º). The Matthews coefficient suggests one molecule in the asymmetric unit and a solvent content of 49 %.

Initially the structure determination of the Rm1ClC cytoplasmic domain was attempted by molecular replacement. The known structures of three CBS domains of eukaryotic ClC proteins (ClC-0, ClC-5 and ClC-Ka) and one structure of the prokaryotic

homolog PtClC provide different possible search models. Because of the similar size to the Rm1ClC domain and the similar crystallization characteristics (1 molecule in AU, structure determined at 1.8 Å) the structure of the PtClC domain appeared to be the most promising search model.

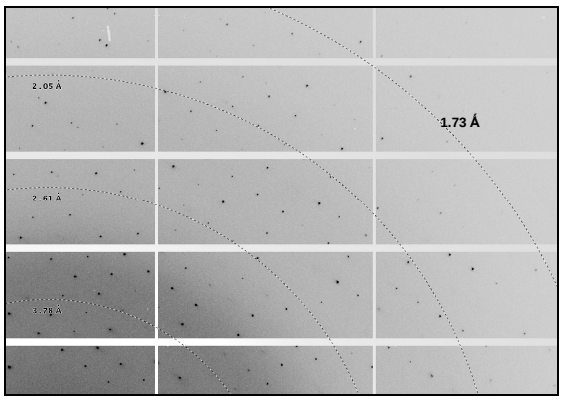



**Figure 2.3.4. Diffraction image of Rm1ClC cytoplasmic domain crystal.**
Crystals grown in 200 mM $CaCl_2$, 40 % PEG, pH 9.4 diffracted to 1.7 Å. Resolution rings are indicated. The data was collected on a frozen crystal on the X06SA beamline at the SLS on a PILATUS detector.

Unfortunately, all molecular replacement trials failed. The reason for this failure, which was previously also encountered for the structure determination of other ClC cytoplasmic domains, might be the large variation in the mutual arrangement of the two CBS subdomains. Independent searches with either of the two subdomains (CBS1 or CBS2) of different proteins as search models were also unsuccessful. Therefore, the structure determination was pursued by experimental phasing.

### 2.2.3.4 Structure determination of the Rm1ClC domain

Experimental phases can either be determined by isomorphous replacement or anomalous scattering methods that rely on information that can be derived from small differences within a dataset (for anomalous scattering) or between diffraction datasets (for isomorphous replacement). In both methods, the first step is the location of the heavy atoms or anomalous scatterers, generally termed the substructure, in the crystallographic asymmetric unit.

In order to determine the structure by either isomorphous replacement or anomalous scattering, the crystals of the Rm1ClC cytoplasmic domain were derivatized with different heavy-atoms and ions. Crystals were soaked in solutions containing low mM amounts of different salts of heavy atoms like Pt, Hg, Pb, Au, Eu, W and higher concentrations of the halides $I^-$ and $Br^-$. Initial trials to solve the structure by single or multiple isomorphous replacement were unsuccessful. Although, the binding of mercury

and iodide was detected by manual inspection of difference Patterson maps (Figure 2.3.5), the derivatized crystals were not isomorphous with the native crystals, which prevented structure determination. The frequently occurring problem of non-isomorphism is avoided when using anomalous dispersion methods like single (SAD) and multiple wavelength anomalous dispersion (MAD).

The structure of the Rm1ClC cytoplasmic domain was determined by SAD using the anomalous scattering of mercury in the crystals derivatized by mercury (II)-chloride. The data was collected at the L-III X-ray absorption edge of mercury for maximizing anomalous differences. The data was collected with high redundancy and an attenuated beam was used. The high redundancy increases precision of the data and attenuated beam reduces radiation damage, both improving the quality of the anomalous signal.

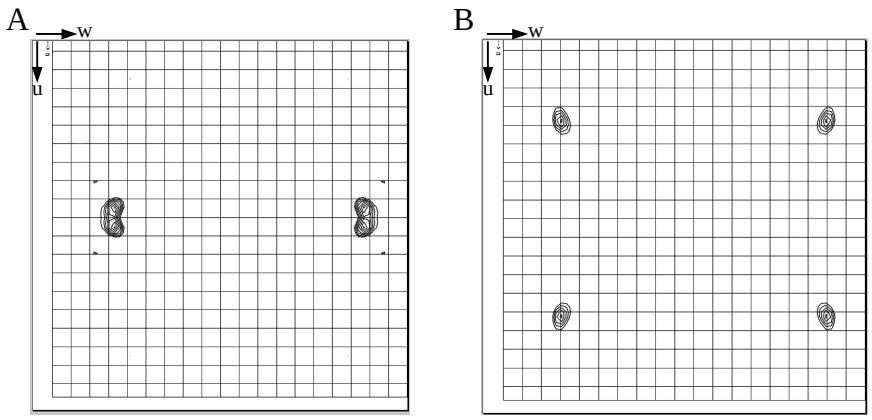



**Figure 2.3.5. Isomorphous difference Patterson maps of Rm1ClC domain crystals derivatized with $HgCl_2$ (A) and $CaI_2$ (B).** Peaks of the mercury and iodide sites are found in the Harker plane u [1/2] w. Contour level starts at 3σ in steps of 1σ.

After scaling and merging of the data (Friedel pairs were treated as independent reflections) the magnitude of the anomalous contribution was analyzed by plotting the anomalous difference ($|F^+-F^-|/\sigma$) against resolution (Figure 2.3.6). As a cut off for phase determination a resolution of 2.2 Å was chosen, where the anomalous difference was 1.3 times over the background.

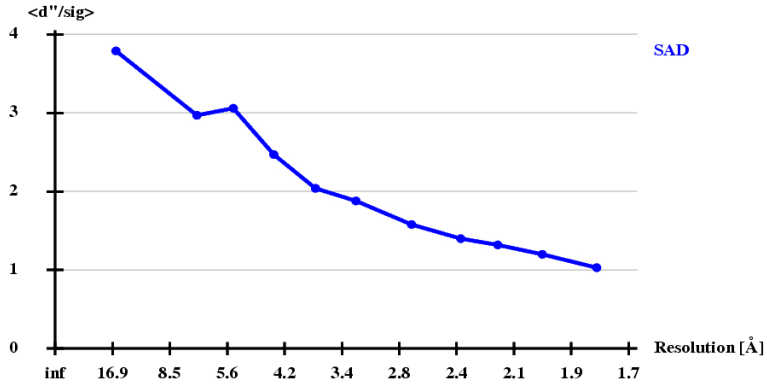



**Figure 2.3.6. Dependence of anomalous signal on the resolution in the SAD data set of Hg-derivatized Rm1ClC domain crystals.** The anomalous difference (d") over the standard deviation of the error of the measurement (σ) is plotted against the resolution. The anomalous signal is detectable to 1.9 Å resolution.

The anomalous difference Patterson map calculated from a Hg-SAD dataset shows strong peaks at Harker sections indicating the binding of mercury to the protein (Figure 2.3.7). Since the Rm1ClC domain contains only one cysteine, one Hg atom in the asymmetric unit was expected, which was easily identified by Patterson interpretation. The mercury coordinates and initial phases were calculated using the program SHELXD (Schneider and Sheldrick 2002).

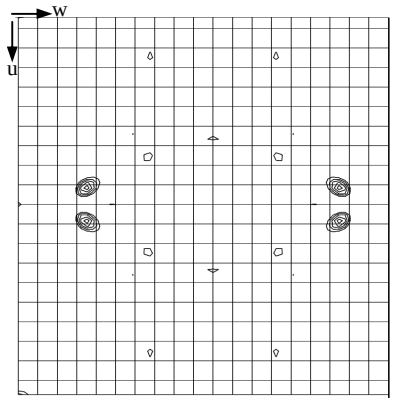



**Figure 2.3.6. Anomalous difference Patterson map of the Rm1ClC domain crystal derivatized with $HgCl_2$.** The data was collected at the L-III X-ray absorption edge of mercury. Strong peaks of the mercury sites are found in the Harker plane u [1/2] w. Contour level starts at 3σ in steps of 1σ.

Initial SAD phases were improved by 20 cycles of density modification and solvent flipping implemented in SHELXE (Sheldrick 2002). During this procedure the correct hand of the mercury coordinates was determined. The experimental electron density map obtained from SHELXE was of very good quality and allowed model building of the largest part of the protein (Figure 2.3.7).

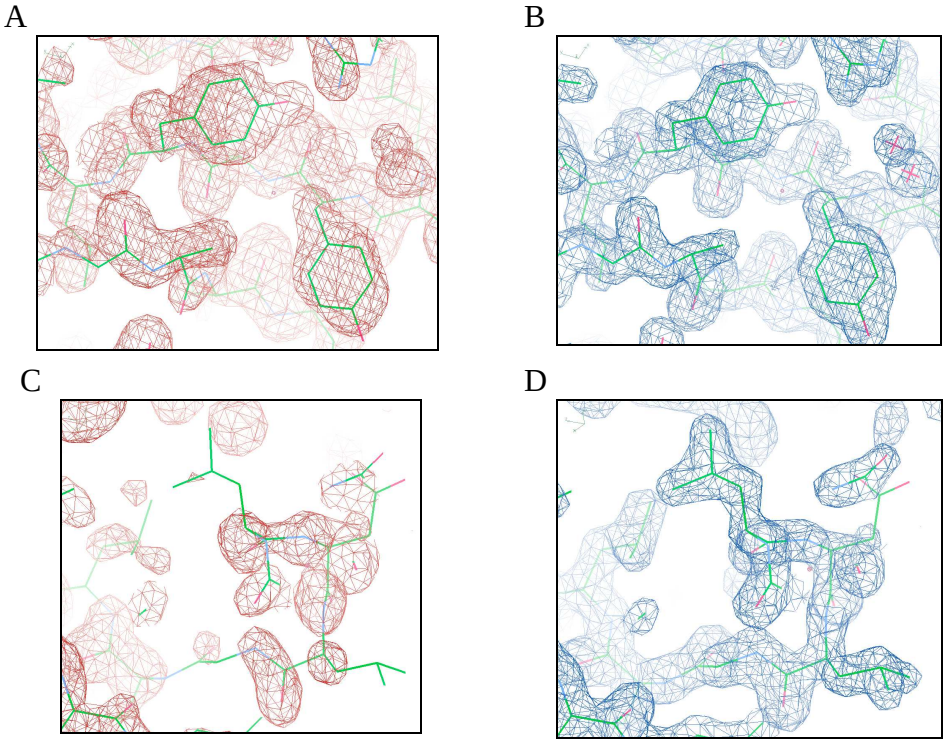



**Figure 2.3.7. Electron density maps of the Rm1ClC cytoplasmic domain.**
**(A) and (C)** Experimental density obtained from SHELXE and used for model building; (**B) and (D)** 2Fo-Fc electron density after completed refinement. The density was calculated at 2.2 Å and 1.6 Å, respectively and contoured at 1.5 σ. The electron density is superimposed on the refined model of the Rm1ClC domain. The figures **A** and **B** represent the section of the electron density in the interface between the two CBS subdomains that was very well defined already in the experimental map. In the linker region (residues 501-505) the electron density in the experimental map was missing (**C**), but after several rounds of refinement the difference map allowed model building in this region too (**D**).

An initial model was built automatically with the program ARP/wARP (Langer, Cohen et al. 2008). From 134 residues, 79 residues were correctly placed in that way. The initial model was manually corrected and partially completed using the graphical program Coot (Emsley and Cowtan 2004). The corrected model was still missing 13 residues in the linker region between the CBS1 and CBS2 subdomain, because the electron density in this region of the experimental map was not observed. After several runs of refinement, the density for the linker region was observed in 2Fo-Fc maps, which allowed building of the missing residues. The final model consists of a total of 130 amino acids including the residues 436–560 of Rm1ClC, a methionine and a serine introduced at the N-terminus by FX cloning and three residues of the 3C protease cleavage site at the C-terminus.

| | Native | Hg |
|---|---|---|
| **Data collection** | | |
| Space group | $P2_12_12$ | $P2_12_12$ |
| Cell dimensions | | |
| a, b, c (Å) | 40.51 88.18 39.13 | 41.41 87.81 39.02 |
| α, β, γ (°) | 90.0, 90.0, 90.0 | 90.0, 90.0, 90.0 |
| Resolution (Å)* | 50-1.6 (1.7-1.6) | 50-1.8 (1.9-1.8) |
| Wavelength (Å) | 1.0707 | 1.006 |
| $R_{merge}$ (%) | 5.6 (48.9) | 13.4 (68.9) |
| $R_{mrgd-F}$ (%) | 6.4 (50.6) | 7.4 (30.2) |
| I/σI | 19.61 (3.12) | 16.75 (4.26) |
| Completeness (%) | 96.8 (86.4) | 97.6 (90.1) |
| Redundancy | 6.4 | 11 |
| **Refinement** | | |
| Resolution (Å) | 19.74 -1.6 | |
| No. reflections | 19086 | |
| Rwork/Rfree (%) | 16.5 / 20.4 | |
| No.atoms | | |
| Protein | 965 | |
| Ion | 1 | |
| Water | 121 | |
| Average B-factors | Isotropic Anisotropic | |
| Protein | 22.38 57 | |
| Ion | 21.63 - | |
| Water | 36.61 - | |
| R.m.s. deviations | | |
| Bond lengths (Å) | 0.006 | |
| Bond angles (%) | 1.002 | |

* Values in the parenthesis are for the highest resolution shell.

$R_{merge} = \Sigma|I(h,i)-I(h)| / \Sigma(I(h,i)$ , where I(h,i) is the mean intensity of the reflection

$R_{work} = \Sigma||Fo|-|Fc|| / \Sigma|Fo|$, where Fo and Fc are the observed and calculated structure factor amplitudes, respectively.

$R_{free}$ was calculated using a randomly selected 5 % sample of the reflection data omitted from the refinement.

**Table 2.3.1. Data collection and model refinement statistics.**

Refinement was carried out against the native dataset to the resolution of 1.6 Å with the program PHENIX.refine (Adams, Grosse-Kunstleve et al. 2002). The first refinement steps included restrained refinement of atomic positions, occupancies and stereochemistry, isotropic refinement of atomic displacement parameters (B factors) and bulk solvent correction. In the following steps, water molecules were included and TLS refinement was performed by dividing the molecule into two groups that move as rigid bodies. Between the refinement steps, the model was manually adjusted to fit better into the calculated difference map (2Fo-Fc). The last refinement step consisted of the anisotropic refinement of B factors of protein atoms as well as of the isotropic refinement of water molecules and ions. The final model, containing 965 protein atoms, 121 water molecules and 1 chloride ion, agreed with the crystallographic data with $R_{work}$ of 16.5 % and $R_{free}$ of 20.4 % and did not show outliers in Ramachandran plot.

### 2.2.3.5 Structure of the Rm1ClC cytoplasmic domain

The structure of the Rm1ClC cytoplasmic domain is presented in Figure 2.3.8. The overall organization resembles the equivalent structures of the cytoplasmic domains of ClC-0, ClC-5, ClC-Ka and PtClC. The domain consists of two subdomains, CBS1 and CBS2, which are connected by a short linker. The two subdomains are related by a pseudo-two-fold arrangement and tightly interact via an interface formed by a pair of β strands.

The secondary structure elements in the CBS motifs of Rm1ClC are slightly different than found in eukaryotic ClC members (β1, α1, β2, β3, α2). In both CBS1 and CBS2 the short β1 strands are absent, hence giving the motive α1, β1', β2', α2 (Figure 2.3.8.D). The absence of β1 strands has already been observed in the CBS subdomains of PtClC, suggesting that this could be a general feature of the cytoplasmic domains in prokaryotic ClC proteins.

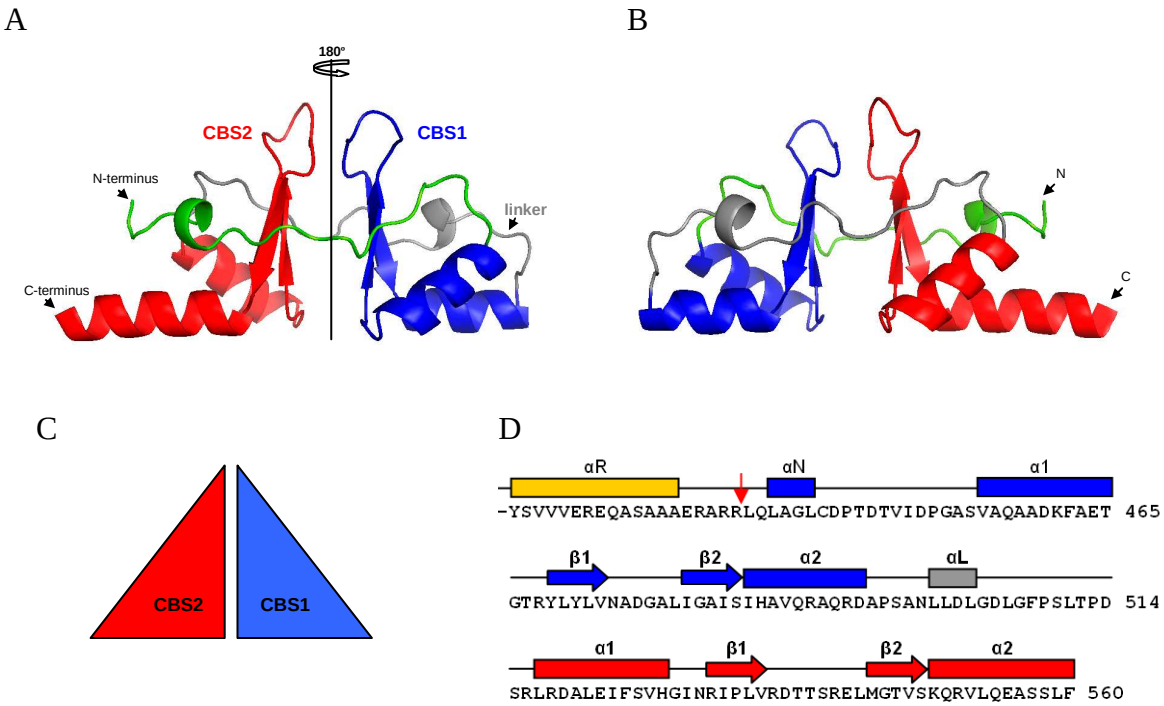



**Figure 2.3.8. Structure of the cytoplasmic domain of Rm1ClC.**
(A) Ribbon representation of the Rm1ClC domain. The CBS1 subdomain is colored in blue, the CBS2 in red. The N-terminal region connecting the CBS1 with the R helix is colored in green and the linker region between to CBS subdomain in gray. The model in (A) is rotated by 180º around the pseudo-two-fold axis to give the model orientation in (B).
(C) Schematic representation of Rm1ClC domain in the same orientation as in (A).
(D) Sequence of the Rm1ClC CBS domain with its secondary structure elements indicated. The R helix is colored in yellow, the elements belonging to CBS1 in blue, the helix in the linker region in grey and the elements of CBS2 in red. The red arrow is indicating the first residue used for cloning and crystallization.

The N-terminal region that connects CBS1 with the transmembrane domain is 19 residues long and consists of a short helix (αN) followed by a long coil that wraps around the protein. The linker connecting the two CBS subdomains consists of 22 residues and tightly interacts with the residues from both subdomains. This is probably the reason for its defined electron density, which was not the case for the longer linkers found in the eukaryotic CBS domains. The N-terminal coil and the linker are located at about the same height on the opposite sides of the cytoplasmic domain surface and encircle the β sheets of the two CBS subdomains in different directions. The Rm1ClC CBS domain does not contain a C-peptide as the last amino acid is part of the CBS2 α2 helix.

Besides its small size and compactness, another interesting feature found in the Rm1ClC domain structure is a very symmetric shape with respect to the secondary structure elements in CBS1 and CBS2 (Figure 2.3.9). The pseudo-two-fold arrangement extends from the two β sheets to other parts of the subdomains. The reason for this symmetry is the same length of the corresponding α helices and β strands in the two subdomains. The structural conservation between the two subdomains is also illustrated by the low rmsd value of 1.07 Å obtained when superimposing the Cα positions of both subdomains. This feature was not observed to this extent in other crystallized ClC CBS domains.

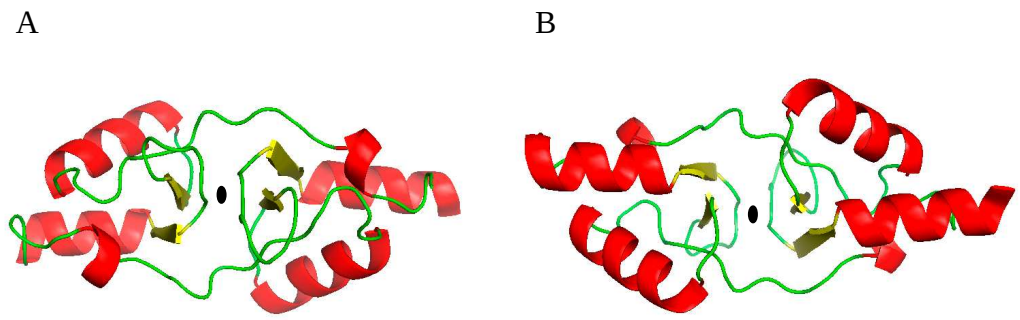



**Figure 2.3.9. Pseudo-two-fold symmetry of the Rm1ClC cytoplasmic domain.** (**A**) A view from the bottom, (**B**) a view from the top. The helices are colored in red, the strands in yellow, the loops and coils in green. The black dot indicates the centre of symmetry.

The overall structure of the Rm1ClC domain was compared to the other ClC cytoplasmic domains by superimposing their main chains. Using the SSM (secondary structure match) algorithm (Krissinel and Henrick 2004), the highest similarity was found with the domain of PtClC with a root mean square deviation (rmsd) of 2.28 Å. Superposition with the ClC-0, ClC-5 and ClC-Ka gave rmsd values of 2.49, 2.51 and 2.78 Å, respectively.

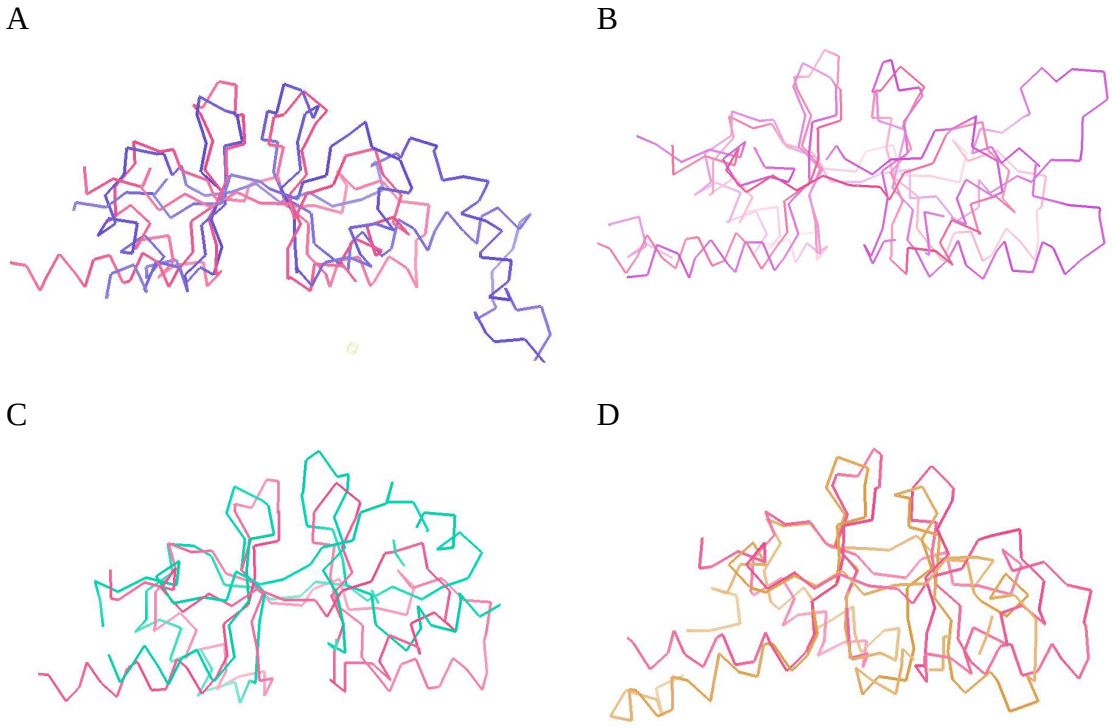



**Figure 2.3.10. Structural comparison of ClC cytoplasmic domain.**
The structure of the cytoplasmic domain of Rm1ClC (red) is superimposed with the cytoplasmic domains of ClC-0 (A, blue); ClC-5 (B, magenta); ClC-Ka (C, green) and PtClC (D, orange). The superposition was generated by a least-square fit of the respective CBS2 subdomains. The angle between CBS1 and CBS2 is similar to the one of ClC-0 and ClC-5, but different than found for ClC-Ka and PtClC.

When superimposing the cytoplasmic domains by a least-square fit of the CBS2 subdomains, differences in the arrangement between CBS1 and CBS2 can be observed (Figure 2.3.10). The angle between the β sheets in the Rm1ClC domain is similar as found in the domains of ClC-0 and ClC-5, while it is smaller than the equivalent angle in the ClC-Ka domain and bigger than in the PtClC domain. The cleft in the interface of the two CBS subdomains is the place where nucleotide binding was detected in the ClC-5 domain. Although the overall size of the cleft in Rm1ClC appears similar to ClC-5 when comparing the main chain atoms only, it is much smaller when including side chains. The residues found in this region are not conserved between the two proteins and the interactions with nucleotide binding would not be possible in Rm1ClC for sterical reasons. However, it is still possible that other ligands bind to this region of Rm1ClC, whose identification will require extensive biochemical screening.

### 2.2.3.6 Ion binding to the Rm1ClC cytoplasmic domain

Besides water molecules, a putative single chloride ion was detected in the electron density of the Rm1ClC cytoplasmic domain structure. The binding of ions to the CBS domains of other protein families has been previously observed (Biemans-Oldehinkel, Mahmood et al. 2006; Hattori, Iwase et al. 2009) In order to systematically investigate ion binding to the Rm1ClC domain, the crystals were soaked in solutions containing different anions and cations.

Initially, the binding of halides was tested. The crystals were incubated in solutions containing $Br^-$ and $I^-$ and data allowing the calculation of anomalous or isomorphous differences were collected. Interestingly, the binding of both bromide and iodide was detected at the same position where the putative chloride ion was bound in the native crystals (Figure 2.3.11A and B). Two arginine residues (R516 and R518) surround the halide ion and provide with their positively charged side chains a suitable chemical environment for anion coordination. To quantify the affinity of chloride binding to the isolated Rm1ClC cytoplasmic domain, biochemical experiments like isothermal titration calorimetry or equilibrium dialysis need to be performed. It will be also important to reveal whether this binding site also exists in the context of the full length Rm1ClC protein.

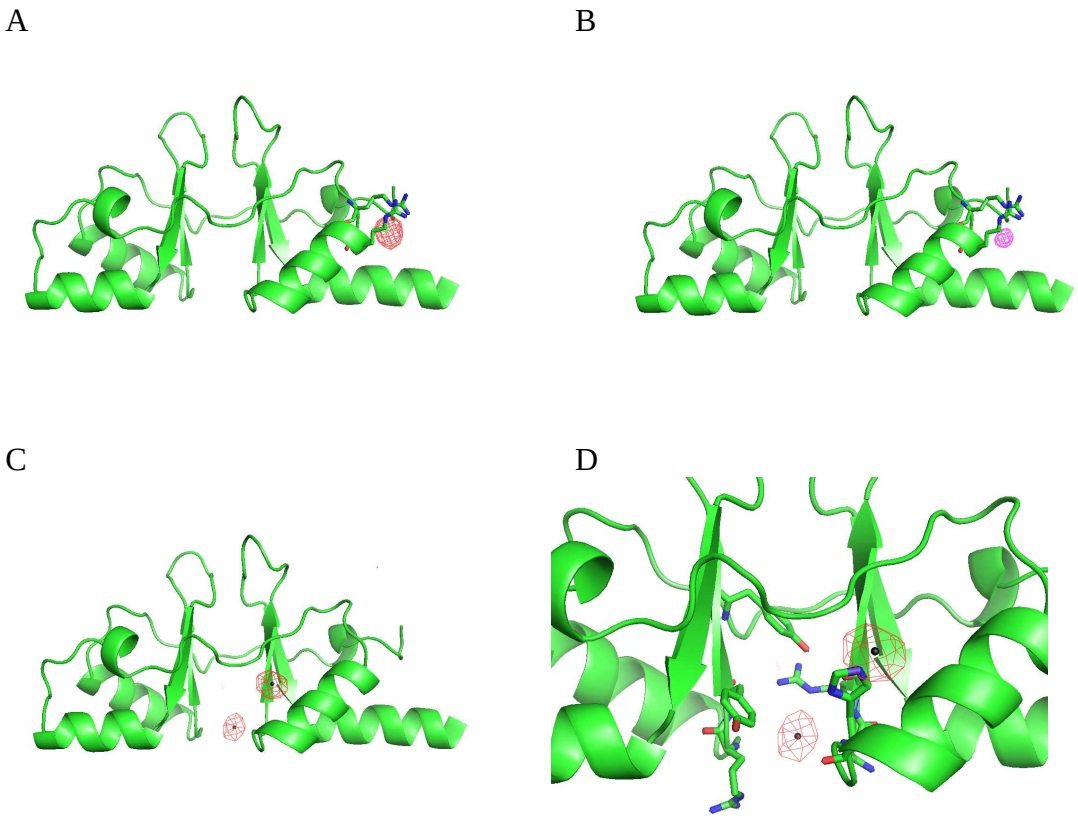



**Figure 2.3.11. Ion binding to the Rm1ClC cytoplasmic domain.**

**Figure 2.3.11. Ion binding to the Rm1ClC cytoplasmic domain.**
Crystals of the Rm1ClC domain were soaked in solutions containing different ions and the ion binding was crystallographically detected by calculating difference maps. (A) Electron density of bromide found in the anomalous difference map of a crystal soaked in 100 mM NaBr. The map is contoured at 3 σ. (B) Electron density of iodide found in the isomorphous difference map of a crystal soaked in 400 mM $CaI_2$. The map is contoured at 8 σ. (C) Electron density of $Zn^{2+}$ ions found in anomalous difference map of a crystal soaked in 100 mM $ZnCl_2$. The map is contoured at 4 σ. (D) Closer view on the residues in the region of bound $Zn^{2+}$ ions.

The binding of different cations to the Rm1ClC domain was tested in a similar way. By soaking the crystal in solution containing the ion the specific binding of two $Zn^{2+}$ ions was detected (Figure 2.3.11.C and D). One of them is located exactly in the middle of the cleft between the two CBS subdomains. The ion is coordinated by interaction with the side chains of two arginines (R467 and R531) and one asparagine (N530). The second $Zn^{2+}$ ion binds at the end of the helix α1 of CBS2 and it is coordinated by a histidine residue (H527). Since the ions were found in crystals grown at pH 9.4, it would be clearly interesting to study the pH dependence of binding in order to see whether the $Zn^{2+}$ ion would also be bound at physiological pH.

Rm1ClC originates from a bacterium that grows in heavy metal polluted water (von Rozycki, Nies et al. 2005). This microorganism can survive millimolar concentrations of $Zn^{2+}$, $Cd^{2+}$, $Co^{2+}$, $Pb^{2+}$, $Hg^{2+}$, $Ni^{2+}$ and $Cr^{3+}$. The identification of $Zn^{2+}$ ions bound to the cytoplasmic domain of Rm1ClC thus opens the question on the potential physiological importance of heavy metal binding. Heavy metals might play a role of ligands for the cytoplasmic domain of this ClC protein and in this way influence ion transport. Future experiments will include investigations on the binding of other heavy metals by crystallography and biochemical and functional assays.

### 2.2.3.7 Possible dimeric arrangement of the Rm1ClC cytoplasmic domain

The asymmetric unit of Rm1ClC domain crystals contains only a single subunit. By sedimentation analysis it was also shown that this domain exists as a monomer in solution. Both results suggest that there are no strong interactions between the isolated cytoplasmic domains. The question whether the cytoplasmic domains interact in the context of the full length dimer is difficult to answer in the absence of a full-length protein structure, but it is still possible that the relative high concentration of the domains in the full length protein allows weak interactions to be formed. Possible interaction modes for an Rm1ClC domain dimer can be observed in the crystal by analyzing the contacts between symmetry related molecules and by analogy to the observed dimeric arrangement of eukaryotic ClC cytoplasmic domains.

Since the crystals of the Rm1ClC domain were of the space group $p2_12_12$, only one 2-fold relationship was found along the z axis. Figure 2.3.12A shows this arrangement. The interaction is formed between the CBS2 subdomains. The interacting residues are from the helices α1 and α2 and the combined molecular surface buried in the interface is 1022 $Å^2$, as calculated in the PISA program (Krissinel and Henrick 2007). Five hydrogen bonds between interacting residues stabilize the interaction. In this dimeric assembly both the N- and the C-termini of the domain, and the bound chloride ions are located close to the interface while five residues make the connection to the predicted R-helix.

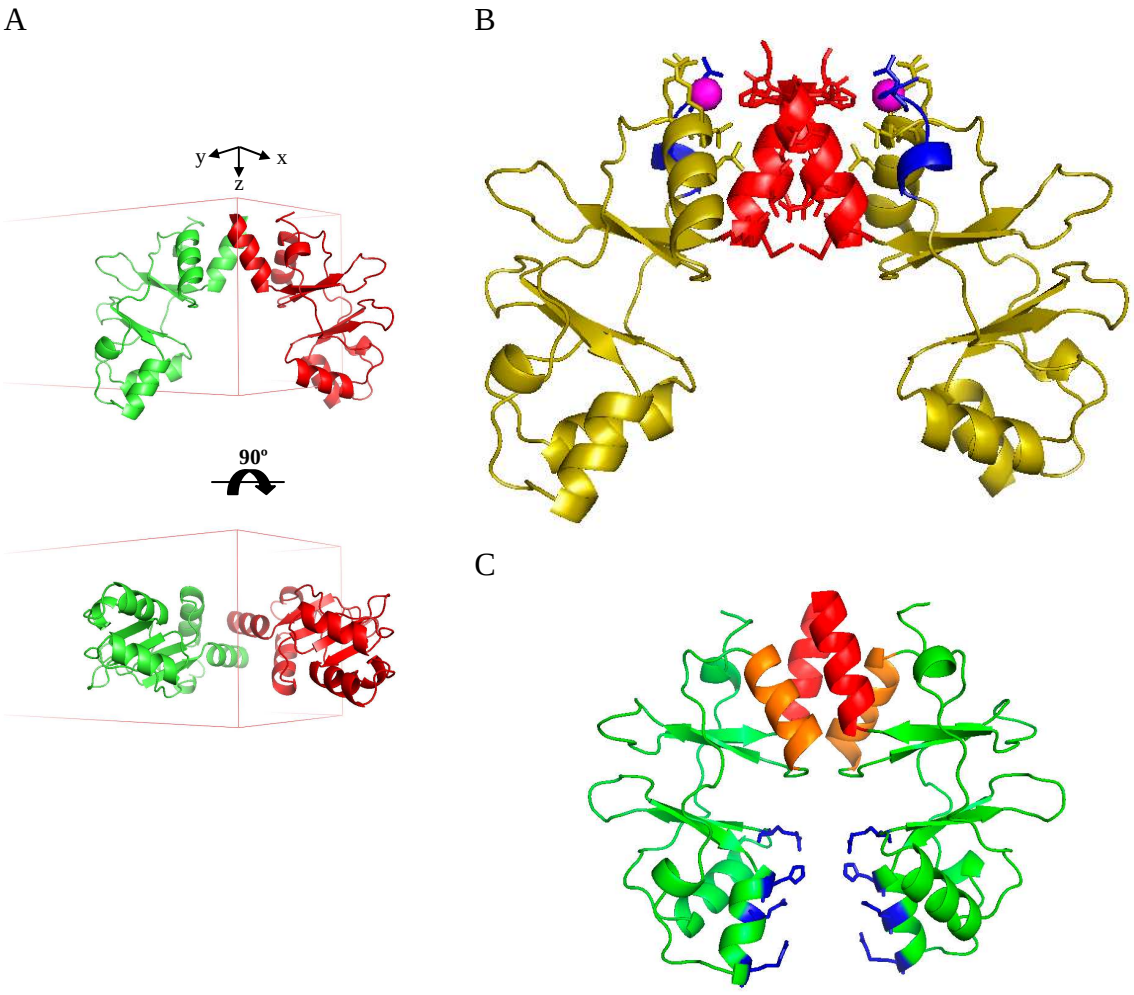



**Figure 2.3.12. Two-fold relationship in the crystal of the Rm1ClC domain.**
**(A)** Dimer related by crystallographic symmetry viewed from the side (top) and from the top (bottom).
**(B)** Close-up of A. Residues in the interface are represented as sticks, the N-terminus is colored in blue, the C-terminus in red and the bound chloride ions are represented as magenta spheres.
(C) Dimer of the Rm1ClC domain created by superposition on the TM0892 dimer (PDB entry 1VR9). Helices α2 are colored in red, helices α1 in orange. The positively charged residues from the CBS1 are colored in blue.

The observed dimeric arrangement in the crystal of the Rm1ClC domain looks similar to the head-to-head dimer of a CBS protein from *Thermatoga maritima* TM0892.

The interaction in the TM0892 dimer is mainly formed by the residues in the helices α1 and α2 of the CBS2 subdomain, with the minor contacts between the residues from the α2 of the CBS1 subdomains. This interface has a combined surface of 1420 Å$^2$, and is quite hydrophobic. A hypothetical dimer of the Rm1ClC based on the superimposition of CBS2 subdomains of TM0892 and Rm1ClC (Figure 2.3.12C) contains several clashes between the residues in the helices of CBS2. It would also place positively charged residues in the α2 of the CBS1 into proximity that might lead to repulsions and therefore favor more open structural arrangement as found in the crystal. In order to test whether the interaction of Rm1ClC domain found in the crystal exists in the context of the full length protein, the residues in the interface will ultimately have to be mutated to cysteine and the cross-linking analysis will have to be performed.

The dimeric arrangement observed in the crystals of the cytoplasmic domain of ClC-Ka has been shown to be also conserved in ClC-5 and ClC-0 (see next chapter). As the investigated domains of prokaryotic ClC homologs did not show a similar organization, it would be interesting to reveal whether the observed interactions would be structurally feasible. In order to answer this question, a dimer of the Rm1ClC cytoplasmic domain was constructed that resembles the interaction in ClC-Ka (Figure 2.3.13).

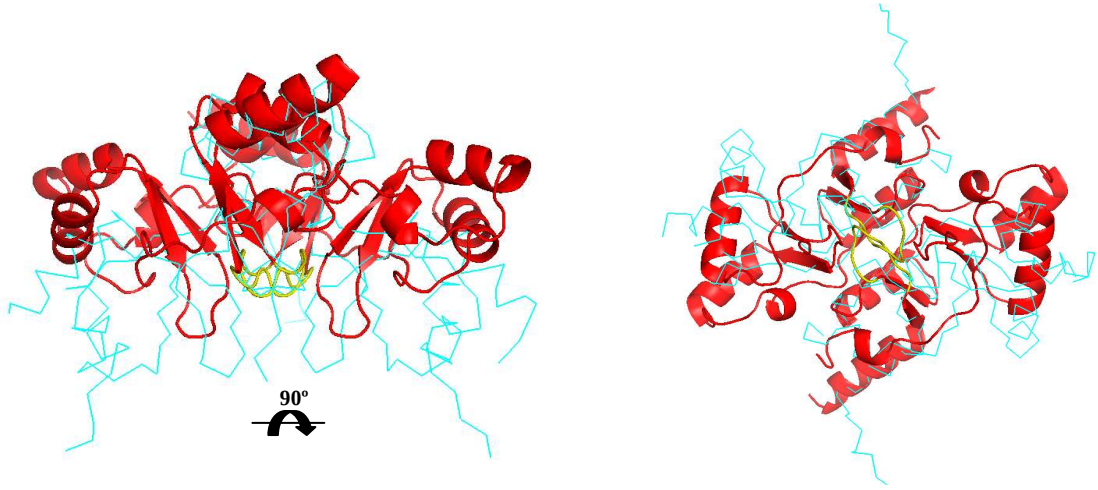



**Figure 2.3.13. Hypothetical dimeric arrangement of the Rm1ClC domain based on the structure of eukaryotic homologs.** The dimer of Rm1ClC domain is colored in red, with the loops connecting β strands in CBS2 in yellow. The ClC-Ka domain dimer is colored in blue. The relationship between the two views is represented.

The domain of Rm1ClC was superimposed on the dimer of the ClC-Ka domain by a least-square fit of the region of CBS2 that contains the interaction interface. The resulting Rm1ClC oligomer looks different from its eukaryotic counterpart. While the ClC-Ka dimer has a V-shape, the Rm1ClC dimer is more compact due to the smaller inclination between the two CBS subdomains. Several clashes between the side chains in the interface are

observed, with the most prominent ones between the loops connecting two β strands in CBS2. Like other potential interactions described before, this hypothetical arrangement thus needs to be investigated by a combination of mutagenesis and cysteine cross-linking.

### 2.2.3.8 Summary

The expression of isolated cytoplasmic domain of the Rm1ClC protein allowed its crystallization, structure determination and biochemical characterization. Several important features of this domain were revealed including the elucidation of its detailed molecular architecture, the monomeric state in solution and binding of ions to specific sites of the domain. The obtained structural information on the Rm1ClC domain now enables the design of detailed structural and functional experiments that would lead to the better understanding of the role of cytoplasmic domains in prokaryotic ClC proteins.

### 2.2.3.9 Material and methods

**Expression and purification of the Rm1ClC cytoplasmic domain.** The cytoplasmic domain of Rm1ClC encompassing residues 436-560 was constructed by FX cloning into pBXC3C vector and transformed into *E. coli* MC1061 cells. Expression cultures were grown in TB medium supplemented with 100 µg/ml ampicilin. Cells were grown at 37ºC until an $OD_{600}$ of 0.6 and subsequently cooled to 25ºC. Protein expression was induced with 0.002 % arabinose for 14 - 15 hours at 25ºC.

Cells were harvested by centrifugation at 4000 rpm (rotor Sorvall H6000A) for 20 min at 4ºC. The cell pellet was resuspended in the buffer containing 50 mM Tris-HCl pH 7.4 and 150 mM NaCl (buffer C) with addition of 1 mg/ml lysozyme, 20 µg/ml DNase, 1µg/ml leupeptin, 1 µg/ml pepstatin and 1 mM PMSF. The cells were lysed with an Emulsiflex C3 high pressure homogenizer (Avestin) with at least two passages at 700 bar. The lysate was cleared by low-spin centrifugation at 11,500 × g for 40 min at 4ºC and the supernatant was used for protein purification.

All purification steps were carried out at 4ºC. The protein was purified by affinity chromatography on Ni-NTA agarose (Qiagen). After binding of the protein the resin was washed with buffer C until baseline of $OD_{280}$ was reached. To remove non-specifically bound proteins, the resin was washed with the buffer C supplemented with 50 mM imidazole. The protein was eluted with buffer C containing 300 mM imidazole.

To cleave off the $His_{10}$-tag the protein was digested for 2 h with 3C protease at the ratio 1 mg of protease for 10 mg of fusion protein. During digestion the solution was

dialyzed against the buffer C. After cleavage, the solution was incubated for 15 min with Ni-NTA agarose (Qiagen) to remove the His-tagged 3C protease. The cleaved protein was concentrated using an Amicon Ultra Centrifugal filter device (MWCO = 5 kDa, Milipore) to a volume of ~ 400 µl. The concentrated sample was subjected to size exclusion chromatography on Superdex S200 column (GE Healthcare) equilibrated in buffer containing 10 mM Tris-HCl pH 7.4 and 150 mM NaCl.

**Analytical ultracentrifugation.** Sedimentation velocity experiments were performed on a Beckmann model XL-I analytical ultracentrifuge with a 50-Ti rotor at 4ºC. Double-sector counterpieces were filled with 400 µl of protein samples and 420 µl of the buffer used for SEC. Data was acquired at 280 nm in continuous scan mode in 0.003 cm intervals at a rotor speed of 40,000 rpm. Sedimentation analysis of the Rm1ClC cytoplasmic domain was performed at protein concentration of 35, 70 and 150 µM. Data analysis was performed with the c(S) module of Sedfit (Schuck 2000; Schuck, Perugini et al. 2002). The buffer parameters, partial-specific volume of the protein and the corrected sedimentation coefficient $S^{o}_{20,W}$ were calculated by using Sednterp. For data used in this analysis the root-mean-square deviation of all the fits was between 0.008 and 0.01.

**Crystallization of the Rm1ClC cytoplasmic domain.** Crystals of the Rm1ClC domain were obtained at 4 ºC and 20 ºC by the sitting drop vapor diffusion method. Purified protein with the concentration 5-10 mg/ml was mixed in a 1:1 ratio with the reservoir solution containing 40-50 % PEG400, 200 mM $CaCl_2$, 50 mM Glycine pH 9.4 or 50 mM HEPES pH 7.4. Crystals grew after 4-7 days in the space group $p2_12_12$.

Heavy metal derivatives were obtained by soaking native crystals for 3h in mother liquor containing 1 mM $HgCl_2$, 400 mM $CaI_2$, 100 KBr or 100 mM $ZnCl_2$.

As the crystals grew in such a high concentration of PEG400, the additional cryoprotection was not needed. Directly from the crystallization drops, the crystals were flash-frozen in liquid propane and stored in liquid nitrogen.

**Crystallography.** Data were collected on frozen crystals on the X06SA beamline at the Swiss Light Source of the Paul Scherer Institute on a PILATUS detector (Dectris). The data were indexed, integrated and scaled with XDS (Kabsch 1993) and further analyzed and processed with CCP4 programs (Bailey 1994).

The structure of the Rm1ClC cytoplasmic domain was determined by the SAD method using the anomalous contribution of mercury in a crystal derivatized by $HgCl_2$. The Hg site was identified with SHELXC and D (Schneider and Sheldrick 2002). The determination of the correct hand of the substructure and density modification was performed with SHELXE (Sheldrick 2002). The initial model was built automatically with

ARP/wARP (Langer, Cohen et al. 2008). The model was completed by manual building of the missing residues in Coot (Emsley and Cowtan 2004). The model was refined against the native data set to a resolution of 1.6 Å with the program PHENIX.refine (Adams, Grosse-Kunstleve et al. 2002).

The specific binding of bromide and zinc ions to the Rm1ClC cytoplasmic domain was detected in anomalous difference electron density maps and binding of iodide in isomorphous difference electron density maps.

The superposition of the Rm1ClC domain with other CBS domains was done in Coot by using either SSM (Krissinel and Henrick 2004) or LSQ algorithm.

## 2.3 Structure of the dimeric assembly of the cytoplasmic domains in eukaryotic ClC proteins

Similar to the previously discussed prokaryotic homologs eukaryotic ClC chloride channels and transporters have a conserved molecular architecture consisting of a catalytic membrane domain and a regulatory cytoplasmic domain. They function as dimers of two identical subunits. The structural arrangement of the membrane domains within the dimer has been revealed by the structure determination of two bacterial homologs (Dutzler, Campbell et al. 2002). Due to the lack of cytoplasmic domains in these homologs, no structural information on these domains could have been obtained. The later work on isolated cytoplasmic domains of eukaryotic proteins ClC-0 and ClC-5 led to the structural and biochemical characterization of the intracellular components of ClC proteins (Meyer and Dutzler 2006; Meyer, Savaresi et al. 2007). Both ClC-0 and ClC-5 domains have a typical CBS fold and are found as dimers in solution. Dimeric assembly for ClC cytoplasmic domains is expected as it extends the symmetry of the transmembrane part. However, the structures of the ClC-0 and the ClC-5 CBS domain have not provided a clear answer on the dimeric arrangement of cytoplasmic domains in ClC proteins. While the ClC-0 domain crystallized as a monomer, the ClC-5 domain showed a dimeric relationship not previously observed for CBS domain-containing proteins (see Introduction 1.3).

Cytoplasmic domains are attributed to have a regulatory role in ion transport of ClC proteins, but the exact mechanisms of regulation are currently not known. In the case of the channel ClC-1 and the transporter ClC-5 the regulation likely depends on the binding of nucleotides, which might implicate conformational differences between the bound and unbound state of the cytoplasmic domains. In the channel ClC-0 the process of slow gating is regulated by the movements in the cytoplasmic domains. The recent results showed that the closure of the slow gate coincides with the dissociation of cytoplasmic domains and a distance change of 20 Å (Bykova, Zhang et al. 2006). It is, therefore, understandable that

structural information on the interaction between the cytoplasmic domains in the context of dimeric ClC proteins is necessary to explain these functional properties. Whereas the most convincing answers to these important questions will ultimately be obtained from the structures of full-length proteins with different conformations of the cytoplasmic domains, the characterization of cytoplasmic domains in isolation provides important information on their dimeric assembly and may provide an initial structural basis for the regulatory processes. This is illustrated by the work on the cytoplasmic domain of the human kidney channel ClC-Ka presented in the following publication.

### 2.3.1 The structure of the cytoplasmic domain of the chloride channel ClC-Ka reveals a conserved interaction interface

**Aim of this study:** The main objective of this study was to determine the dimeric arrangement of the cytoplasmic domains in eukaryotic ClC proteins that was left ambiguous from the previous investigations. The dimeric assembly of the ClC-Ka domain found in the crystal was shown to be preserved in solution by the combination of mutagenesis, chemical cross-linking and analytical ultracentrifugation experiments. The conservation of the observed interaction within the ClC family was also investigated in complementary experiments on the cytoplasmic domain of ClC-0. Finally, the dimeric arrangement was confirmed in the context of the full length ClC proteins.

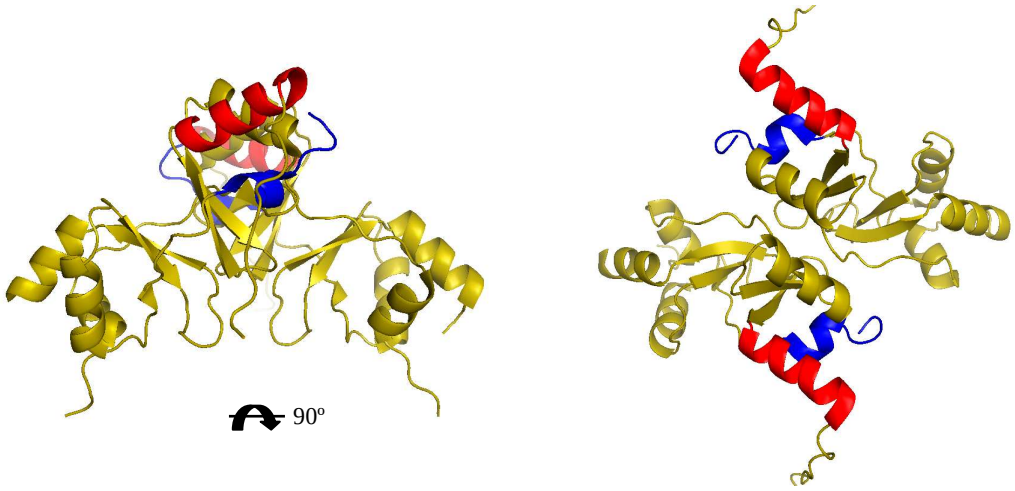



**Figure 2.4.1. Dimeric arrangement of the ClC-Ka CBS domain.**
The protein chains are shown as ribbon, with the N-terminus colored in blue and the C-terminus colored in red. The relationship between the two views is indicated. The V-shaped dimer is formed by the interaction between the CBS2 subdomains.

Structure
Article

# The Structure of the Cytoplasmic Domain of the Chloride Channel ClC-Ka Reveals a Conserved Interaction Interface

Sandra Markovic[1] and Raimund Dutzler[1,*]
[1] Department of Biochemistry, University of Zürich, Winterthurer Strasse 190, CH-8057 Zürich, Switzerland
*Correspondence: dutzler@bioc.unizh.ch
DOI 10.1016/j.str.2007.04.013

## SUMMARY

**The cytoplasmic domains of ClC chloride channels and transporters are ubiquitously found in eukaryotic family members and have been suggested to be involved in the regulation of ion transport. All cytoplasmic ClC domains share a conserved scaffold that contains a pair of CBS motifs. Here we describe the structure of the cytoplasmic component of the human chloride channel ClC-Ka at 1.6 Å resolution. The structure reveals a dimeric organization of the domain that is unusual for CBS motif containing proteins. Using a biochemical approach combining mutagenesis, crosslinking, and analytical ultracentrifugation, we demonstrate that the interaction interface is preserved in solution and that the distantly related channel ClC-0 likely exhibits a similar structural organization. Our results reveal a conserved interaction interface that relates the cytoplasmic domains of ClC proteins and establish a structural relationship that is likely general for this important family of transport proteins.**

## INTRODUCTION

The ClC chloride channels and transporters constitute a large family of membrane proteins that is involved in diverse physiological processes ranging from electrical signaling to epithelial transport and the acidification of intracellular organelles (Jentsch et al., 2005). All eukaryotic family members share a conserved molecular architecture that consists of a complex transmembrane transport domain that is followed by a cytoplasmic domain (Dutzler, 2006; Dutzler et al., 2002). This cytoplasmic component contains a pair of CBS motifs, small protein domains that were named after the enzyme cystathionine β-synthetase (Bateman, 1997). CBS motifs usually occur as tightly interacting pairs and are found to constitute regulatory domains in different protein families such as kinases, membrane transporters, and enzymes (Biemans-Oldehinkel et al., 2006; Ignoul and Eggermont, 2005). For example, in the AMP-dependent protein kinase, a regulatory domain composed of four consecutive CBS motifs binds adenosine nucleotides and is suggested to regulate the catalytic activity of the enzyme in response to changing nucleotide concentrations (Scott et al., 2004; Townley and Shapiro, 2007). Although their role in certain enzymes is well established, the regulatory role within the ClC family is less clear. Different experiments, however, have suggested that in the ClC family, these domains play a key role in the regulation of transport in response to intracellular signals (Bennetts et al., 2005; Maduke et al., 1998; Meyer et al., 2007).

We have previously reported the structures of the cytoplasmic domains of the voltage-dependent chloride channel ClC-0 from *Torpedo marmorata* and of the human transporter ClC-5 (Meyer and Dutzler, 2006; Meyer et al., 2007). Both domain structures allowed the structural characterization of the cytoplasmic components of ClC proteins. The structure of the cytoplasmic domain of ClC-5 additionally revealed insight into its specific interaction with nucleotides, and therefore established the basis of nucleotide recognition by certain ClC family members (Meyer et al., 2007). Two questions that could not be clarified in the previous structures concerned the oligomeric organization of the cytoplasmic domains and whether the specific interaction with nucleotides extends to other members of the family. In order to clarify these two questions and to gain insight into the structural organization of other ClC family members, we have investigated the cytoplasmic domain of the human ClC channel ClC-Ka. The two highly homologous channels ClC-Ka and ClC-Kb are distantly related to the muscle-type ClC channels ClC-0, ClC-1, and ClC-2 (Jentsch et al., 2002). Both channels are expressed in epithelia of the kidney and the inner ear, with mutations causing hereditary renal diseases and deafness (Jentsch et al., 2002). The ClC-K channels have unique regulatory properties and were identified to tightly interact with their β subunit Barttin, a putative transmembrane protein (Estevez et al., 2001).

Here we present the structure of the cytoplasmic domain of the human channel ClC-Ka at 1.6 Å resolution. The crystal structure revealed an interaction mode between two protein chains in a homodimeric molecule that is distinct from interactions seen in CBS motif containing domains from other protein families. A similar organization, however, was previously found in crystals of the cytoplasmic domain of the human transporter ClC-5. The

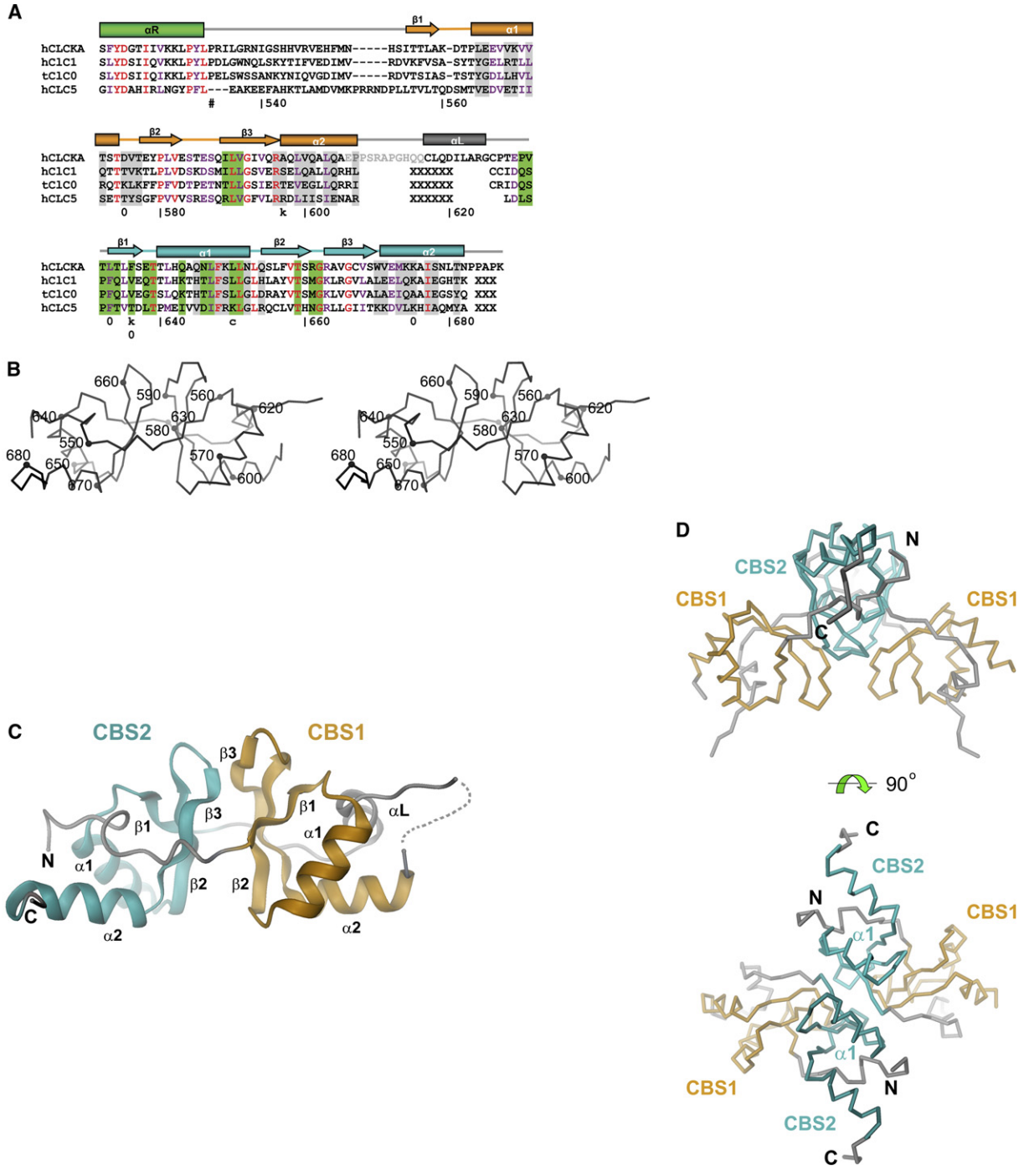



**Figure 1. Structure of the Cytoplasmic Domain of ClC-Ka**

(A) Structure-based sequence alignment of the cytoplasmic domains of the ClC channels ClC-Ka, ClC-1, ClC-0, and the ClC transporter ClC-5. Identical residues are colored in red, similar residues in violet. Residues of the linker of ClC-Ka that lack electron density are colored in gray. Secondary structure and numbering (ClC-Ka) are indicated above and below the sequences, respectively. The linker regions and the C peptides of all regions except ClC-Ka have been omitted (XXX). Residues involved in the observed dimer interface of ClC-Ka are highlighted in green, residues involved in a putative head-to-tail interface are highlighted in gray. The R helix preceding the domain is included in the alignment. The first residue of the crystallized product is highlighted (#). "c" marks the residue in ClC-Ka and ClC-0 used for crosslinking. The residues mutated for analytical ultracentrifugation studies in ClC-Ka and ClC-0 are indicated by "k" and "0," respectively; h, *H. sapiens*; t, *T. marmorata*; hClC-Ka, GenBank accession number P51800; hClC-1, GenBank accession number M97820; tClC-0, GenBank accession number X56758; hClC-5, GenBank accession number 116734718.

(B) Stereo view of a Cα trace of the ClC-Ka domain. Selected residues are labeled according to their position in the hClC-Ka sequence.

high resolution of the data allowed a detailed interpretation of the extended interface, which buries a number of residues that are conserved among different members of the ClC family. In an attempt to probe whether the interaction interface found in the crystal is preserved in solution, we disrupted dimerization by site-directed mutagenesis and introduced specific crosslinks. Both experimental strategies strongly suggest that the crystal structure is representative of the interactions in solution. We subsequently extended our investigations to the cytoplasmic domain of the channel ClC-0. The same approach allowed us to demonstrate that ClC-0, whose oligomerization was not revealed in its crystal structure, shows a similar structural organization. Finally, we extended the crosslinking approach to full-length ClC-0, demonstrating that the same disulfide bond that covalently links the two interacting subunits can also be formed in the context of the transmembrane protein. Both features, the structural similarity found in the dimeric structures of the cytoplasmic domains of ClC-Ka and ClC-5 as well as the similar biochemical behavior of the ClC-Ka and the ClC-0 domain, point at a conserved interaction that is common to eukaryotic ClC family members and that is likely preserved in the full-length proteins where the domains are attached to the transmembrane pore. Our data clarify the previously ambiguous oligomeric structure of the cytoplasmic components of ClC channels and transporters, and reveal important insights into the structural organization of the family.

## RESULTS

### The Structure of the ClC-Ka Domain

To gain insight into the structure of the cytoplasmic domain of the human channel ClC-Ka, we have cloned and expressed a fragment of the protein encompassing residues 533–687 (Figure 1A). The construct includes the entire domain followed by a protease cleavage site and a hexahistidine tag that was cleaved off during purification. This protein construct allowed us to grow crystals of space group $P2_1$ diffracting to a resolution of 1.6 Å. The crystal structure was determined by the multiple isomorphous replacement method; it contains two protein chains, 238 water molecules, one $I^-$ ion, and three $Cl^-$ ions in the asymmetric unit (Table 1). The two protein chains have similar conformations and are related by 2-fold noncrystallographic symmetry. Their structure is well defined in the electron density, with the exception of ten residues of the linker connecting the two CBS domains and ten residues of the N terminus that are disordered.

Figure 1B shows the structure of the ClC-Ka domain. The overall organization resembles the equivalent structures of ClC-0 and ClC-5. Similar to structurally homologous protein domains, the two CBS subdomains (CBS1 and CBS2) are related by a pseudo-two-fold arrangement and tightly interact via an interface formed by a pair of β strands (Figure 1C). The N terminus preceding CBS1 is similar in length and structure to the equivalent region of the channel ClC-0, whereas it is elongated by five residues in the nucleotide binding domain of the transporter ClC-5 (Figure 1A). The short linker region connecting the two CBS subdomains, in contrast, differs from the respective regions of both ClC-0 and ClC-5. Apart from 11 disordered residues following CBS1, the linker forms a short helix (αL) followed by a loop, both interacting with residues of CBS1 (Figures 1A and 1C). To investigate whether the cytoplasmic domain of ClC-Ka binds nucleotides, we collected data from crystals that were grown in the presence of 5 mM ATP (Table 1). An $F_o - F_c$ difference electron density calculated after refinement of the model failed to show any evidence of ATP binding (data not shown). A similar approach revealed the specific interaction of nucleotides with the cytoplasmic domain of ClC-5 (Meyer et al., 2007). When comparing the structure of the nucleotide binding region of ClC-5 with the equivalent structures of ClC-0 and ClC-Ka, it is evident that the shorter N-terminal loop as well as differences in residues in the binding site would probably interfere with nucleotide binding.

### Oligomeric Assembly of the ClC-Ka Domain

Analytical ultracentrifugation experiments of the ClC-Ka C terminus revealed a dimeric organization in solution. A similar sedimentation behavior was previously found for cytoplasmic domains of ClC-0 and ClC-5, thus pointing at a specific interaction of two cytoplasmic domains in the dimeric ClC proteins (Meyer and Dutzler, 2006; Meyer et al., 2007). In accordance with their assembly in solution, a pair of interacting proteins each containing two CBS motifs was found in the asymmetric unit of the ClC-Ka domain crystals (Figure 1D). The high resolution of our data allows a detailed analysis of the protein-protein interface (Figure 2). This structural arrangement mainly involves residues on the surface of CBS2 in each chain of the homodimeric protein. The interactions lead to a V-shaped oligomeric structure which puts the two CBS2 subdomains in close proximity with few mutual interactions between the remote CBS1 counterparts (Figures 1D and 2B). The dimer buries 1530 $Å^2$ of the combined molecular surface and includes more than 25 well-defined water molecules, some of which form a continuous chain bridging a narrow tunnel on one end of the interaction interface (Figures 2A and 2C). Apart from H-bonding interactions between protein residues, which in several cases are mediated by buried water molecules, the bulk of the interactions are hydrophobic. Several residues contributing to this interface are conserved

(C) Ribbon representation of the ClC-Ka domain. The two CBS subdomains are colored in orange and cyan, respectively. Secondary structure elements are indicated.
(D) Cα representation of the ClC-Ka domain dimer in two orientations. The relationship between the two views is indicated. The individual subdomains are labeled and colored in orange and cyan, respectively.
Figures 1, 2, and 5 were prepared with DINO (http://www.dino3d.org).

**Table 1. Data Collection and Model Refinement Statistics**

Data Collection

| | SeMet | Native | MeHg | HgCl | KI | ClC-Ka | ClC-Ka/ATP |
|---|---|---|---|---|---|---|---|
| | Synchrotron | Home | Home | Home | Home | Synchrotron | Synchrotron |
| Space group | $P2_1$ | $P2_1$ | $P2_1$ | $P2_1$ | $P2_1$ | $P2_1$ | $P2_1$ |
| Cell dimensions | a = 35.5, b = 82.0, c = 60.4 | a = 35.2, b = 83.1, c = 59.8 | a = 35.2, b = 81.6, c = 59.9 | a = 35.4, b = 83.6, c = 59.4 | a = 35.2, b = 82.9, c = 59.5 | a = 35.3, b = 83.4, c = 58.9 | a = 35.3, b = 82.5, c = 59.9 |
| β | 93.5 | 92.4 | 92.4 | 91.5 | 92.5 | 92.0 | 92.7 |
| Wavelength (Å) | 0.979 | 1.542 | 1.542 | 1.542 | 1.542 | 1.0 | 1.0 |
| Resolution (Å) | 50–3.0 | 50.0–2.6 | 50–2.8 | 15–3.5 | 15–3.0 | 50–1.6 | 50–1.8 |
| $R_{sym}$ | 11.0 (22.2) | 3.9 (12.4) | 9.9 (37.3) | 14.0 (27.6) | 10.3 (27.6) | 5.5 (27.8) | 6.5 (24.3) |
| I/σI | 10.7 (3.3) | 18.2 (6.0) | 8.4 (2.6) | 7.2 (3.7) | 10.0 (4.6) | 16.9 (2.8) | 17.8 (4.5) |
| Completeness (%) | 93.4 (82.8) | 95.5 (82.1) | 89.8 (86.4) | 90.0 (80.3) | 99.3 (98.9) | 96.9 (79.0) | 98.0 (93.2) |
| Mosaicity (°) | 1.08 | 0.84 | 1.01 | 1.56 | 1.24 | 0.44 | 0.54 |

Refinement

| | Resolution (Å) | Reflections | $R_{work}/R_{free}$ | Atoms | Average B factor | Rmsd Bonds (Å) | Rmsd Angles (°) |
|---|---|---|---|---|---|---|---|
| ClC-Ka | 15–1.6 | 43,605 | 20.0/21.8 | 2,573 | 30.0 | 0.004 | 1.1 |
| ClC-Ka/ATP | 15–1.8 | 31,140 | 20.0/22.7 | 2,573 | 32.9 | 0.004 | 1.2 |

Native was used as a native data set for phasing and initial stages of refinement. $R_{sym} = \Sigma|I_i - \langle I_i \rangle|/\Sigma|I_i|$, where $I_i$ is the scaled intensity of the ith measurement and $\langle I_i \rangle$ is the mean intensity for that reflection. Values for the highest resolution shell are given in parentheses. $R_{work} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where $F_o$ and $F_c$ are the observed and calculated structure factor amplitudes, respectively. $R_{free}$ was calculated using a randomly selected 10% sample of the reflection data omitted from refinement.

between different family members, thus suggesting that this interaction is general also for other ClC channels and transporters (Figures 1A and 2C). Indeed, a similar dimerization mode has previously been identified in the structure of the cytoplasmic domain of the transporter ClC-5, whereas no two-fold relationship was found in the crystals of the corresponding domain of ClC-0 (Meyer and Dutzler, 2006; Meyer et al., 2007). The oligomeric arrangement is unusual for CBS motif containing proteins, which, although frequently found to dimerize, usually interact via an extended interface formed by the four helices α1 and α2 of the two CBS subdomains, in either a "head-to-tail" or "head-to-head" arrangement (Miller et al., 2004; Rudolph et al., 2007; Townley and Shapiro, 2007). The common head-to-tail arrangement, in which the α helices of the subdomain CBS1 from one chain form a four-helix bundle with the equivalent regions in the subdomain CBS2 of the interacting protein, motivated the postulation of a similar interaction mode in ClC-0 (Meyer and Dutzler, 2006).

## Probing the Dimer Interface by Mutagenesis

To clarify whether the assembly in the crystals corresponds to the dimer found in solution, we mutated residues of the dimer interface to either stabilize the oligomeric structure by crosslinking or alternatively to disrupt dimerization. From the structure, we identified the residue Leu 650 to be located at the periphery of the interaction interface in proximity to the two-fold axis (Figures 2B and 2C). The mutation L650C leads to the formation of an intersubunit disulfide bond bridging the introduced cysteine residues under oxidizing conditions (Figures 3A and 3B). The tight structural constraints to form this short covalent bond are fulfilled in the observed dimer, whereas they would be incompatible with a head-to-tail or head-to-head arrangement of the two protein chains.

In a complementary approach, we introduced mutations in residues buried in the dimer interface, aiming to interfere with dimerization. Several mutants involving the conserved leucine residues (Leu 651, Leu 647, Leu 633, and Leu 635) that form the hydrophobic part of the interaction interface failed to express soluble protein, probably due to the fact that mutations of the bulky side chains that also contribute to the hydrophobic core of each protein chain interfere with folding (Figure 2C). Similarly, mutations of other interface residues (such as Thr 634, Thr 639, Thr 659, and Ile 589) did not express soluble protein either (Figure 2C). Mutants involving the interface residue F636, whose side chain is not buried in the protein and instead projects toward the dimer interface, in contrast, did not show any sign of precipitation. The mutant F636D expressed soluble protein with a yield similar to WT. An increase in the gel-filtration elution volume of this mutant already hinted at a smaller molecular weight. The subsequent quantification by analytical ultracentrifugation confirmed that this mutation prevents oligomerization of the domains and that the protein thus sediments as a monomer (Figure 3C). While the mutation of a residue

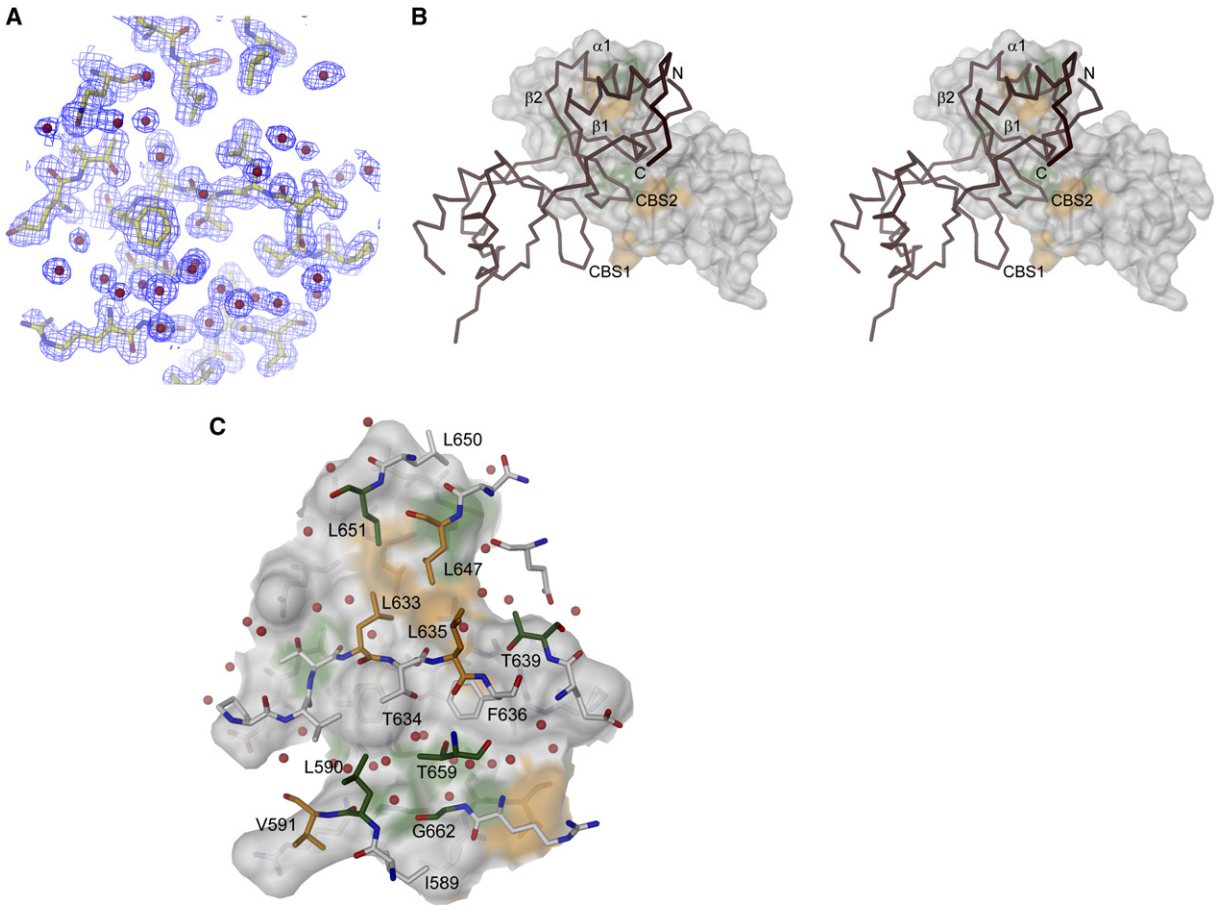



**Figure 2. Structure of the Dimer Interface**

(A) $2F_o - F_c$ electron density of the dimer interface. The density (calculated at 1.6 Å and contoured at 1σ) is shown superimposed on selected residues of the interface. Protein residues are shown as sticks, ordered water molecules as spheres.

(B) Stereo view of the ClC-Ka dimer. The protein is shown as a Cα trace; the molecular surface of one subunit was calculated with MSMS and is shown superimposed (Sanner et al., 1996). The surface of the dimer interface is colored according to the conservation of the underlying residues: green, strictly conserved; orange, homologous.

(C) Interaction interface. The view and the color code is similar to that in (B). Residues contacting the dimer interface are shown as sticks, buried water molecules as red spheres. The surface shows the interaction interface of one subunit.

involved in the observed dimer interface interferes with oligomerization, a mutation that affects a residue that would be involved in a head-to-tail or head-to-head interface did not have any detectable influence on dimer formation. The mutant A597D expresses soluble protein that elutes with a volume similar to WT in gel-filtration experiments and also shows similar sedimentation properties as WT (Figure 3C).

The conclusive correlation between structure and biochemical behavior, as shown for the mutation L650C that stabilizes the dimer by forming an intersubunit disulfide bond and the mutation F636D that prevents dimerization, strongly suggests that the interface found in the crystal structure is preserved in solution and is therefore not an artifact of crystallization.

## Conservation of the Dimer Interface in ClC-0

The strong conservation of residues in the interface and the similarity of the dimeric structures of the domains from ClC-Ka and ClC-5 motivated us to probe whether the same intermolecular interactions found in ClC-Ka would also be preserved in ClC-0. Although the ClC-0 domain robustly dimerizes in solution, its crystal structure did not reveal its oligomeric organization (Meyer and Dutzler, 2006). In analogy to the mutational analysis of the ClC-Ka domain, we mutated equivalent residues in ClC-0. The residue Leu 741 in ClC-0 is the structural equivalent of the ClC-Ka residue Leu 650. As for the ClC-Ka domain, the mutation L741C forms an intersubunit crosslink in the ClC-0 domain (Figures 4A and 4B), thus suggesting that also in ClC-0, this residue is close to the symmetry axis relating two protein chains. Mutations in two positions in the putative dimer interface of the ClC-0 domain express soluble protein and were therefore subjected to sedimentation studies. Mutations of the residues Phe 724 and Val 727, the structural equivalents of the ClC-Ka residues Leu 633 and Phe 636, to arginine both severely weaken the stability of the ClC-0 domain dimer, shifting

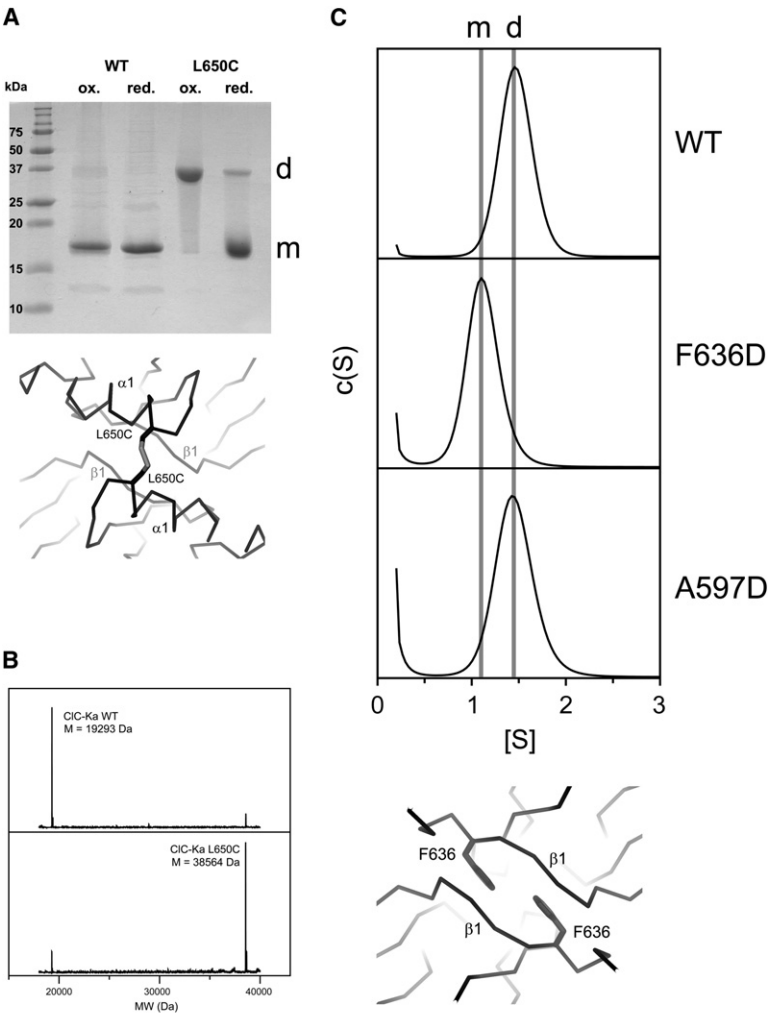



**Figure 3. Crosslinking and Analytical Ultracentrifugation Experiments of ClC-Ka Domain Mutants**

(A) SDS-PAGE of the purified ClC-Ka domain and the domain mutant L650C under oxidizing (ox.) and reducing (red.) conditions. The molecular weight of protein standards and the positions of the monomeric and dimeric proteins are indicated. A model of the crosslinked mutant is shown.

(B) ESI mass spectrometry experiments of the ClC-Ka domain and the mutant L650C. The samples were prepared under oxidizing conditions. The $His_6$ tag was not cleaved off during the experiment. The measurement was within 4 Da of the predicted molecular mass.

(C) Distribution of the sedimentation coefficient (c[S]) as calculated from sedimentation velocity experiments for the WT ClC-Ka domain and the domain mutants F636D and A597D. The oligomeric state (m, monomer; d, dimer) is indicated. A view along the two-fold axis shows the position of the residue Phe 636 in the dimer interface.

the dissociation constant to larger values (Figures 4C and 4D). The mutations K567A and A766R, involving residues that would be part of a possible head-to-tail interface, in contrast, do not have a strong effect on oligomerization in solution, consistent with the assumption that the ClC-0 dimer shares the same interaction interface found in the crystals of the domains of ClC-Ka and ClC-5. Finally, we probed whether a similar crosslink as found for the mutant L741C in the ClC-0 domain could also be formed in the context of the full-length protein. We have expressed full-length ClC-0 and the point mutant L741C as hexahistidine fusion proteins in *Xenopus laevis* oocytes and have investigated the formation of a covalent crosslink of the protein in the membrane under oxidizing conditions following detection by SDS-PAGE and western blot using an anti-$His_6$ antibody (Figure 4E). Also in this case, the mutant L741C allows the formation of a selective crosslink between two protein chains, indicating that the dimeric structure of the isolated domains in solution is preserved in the full-length transmembrane protein.

## DISCUSSION

Ion transport proteins frequently contain cytoplasmic regulatory units that are attached to the transmembrane catalytic domains and that regulate ion transport in response to cellular signals. Such cytoplasmic units are ubiquitously found in eukaryotic members of the ClC family of $Cl^-$ channels and transporters. Although their detailed role in most family members is to date still unknown, increasing experimental evidence points to an involvement in transport regulation, in certain cases in response to the interaction with adenosine nucleotides (Bennetts et al., 2005; Meyer et al., 2007; Vanoye and George, 2002). Two recent crystal structures of the cytoplasmic domains of ClC-0 and ClC-5 revealed the structural organization of these domains from two representative family members (Meyer and Dutzler, 2006; Meyer et al., 2007). However, the oligomeric assembly that relates these putative regulatory components to the dimeric transmembrane pore remained ambiguous. This ambiguity could now be clarified by the structure of the equivalent region of the channel

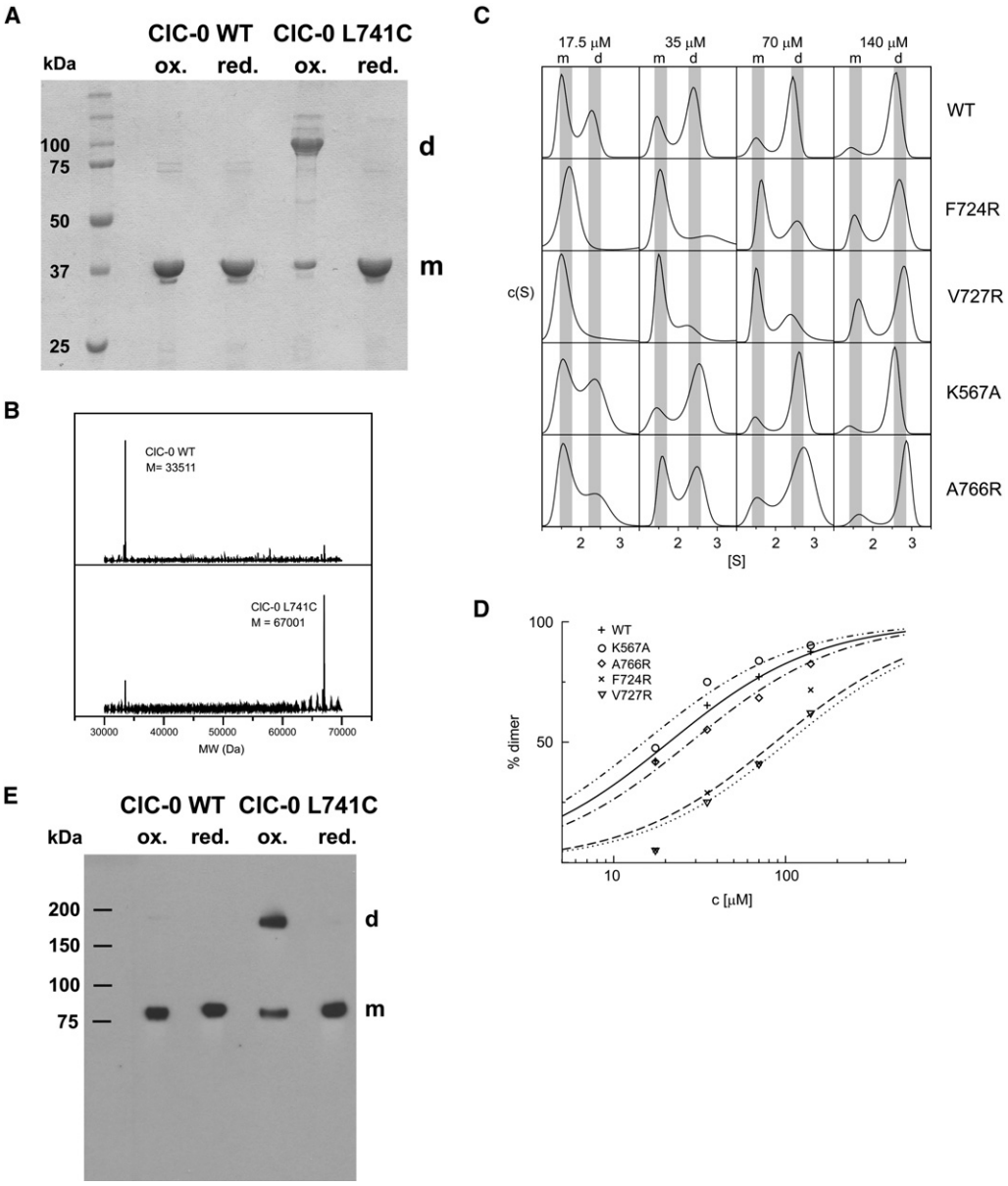



**Figure 4. Oligomerization in the Channel ClC-0**

(A) SDS-PAGE of the purified ClC-0 domain and the domain mutant L741C under oxidizing (ox.) and reducing (red.) conditions. The molecular weight of protein standards and the positions of the monomeric (m) and dimeric (d) proteins are indicated.

(B) ESI mass spectrometry experiments of the ClC-0 domain and the mutant L741C. The samples were prepared under oxidizing conditions. The $His_6$ tag was not cleaved off during the experiment. The measurement was within 7 Da of the predicted molecular mass.

(C) Distribution of the sedimentation coefficient (c[S]) as calculated from sedimentation velocity experiments of the WT ClC-0 domain and of mutants at different protein concentrations. The oligomeric state (m, monomer; d, dimer) is indicated.

(D) Dimer affinity. The fraction of dimeric protein (calculated from the integration of the respective peaks in [C]) is plotted against the protein concentration. $K_d$ was calculated by a fit of the data to the equation $f_{dimer} = 1/(c/K_d + 1)$, where $f_{dimer}$ is the fraction of the dimeric protein and c is the protein concentration. The respective $K_d$s are: WT = 21 μM (—), K567A = 15 μM (-···-), A766R = 28 μM (-··-), F724R = 87 μM (- - -), V727R = 103 μM (···).

(E) Crosslinking of the full-length ClC-0 and the mutant L741C. The protein was expressed as a $His_6$ fusion protein in *X. laevis* oocytes, and crosslinked in its membrane environment under oxidizing conditions. The membrane fraction was separated by SDS-PAGE and detected by western blot using an anti-$His_6$ antibody. The oligomeric states (m, monomer; d, dimer) are indicated. The position of the respective molecular weight markers is indicated on the left (ox., oxidation by incubation in 100 mM sodium tetrathionate at 4°C for 30 min; red., subsequent reduction of the sample after oxidation).

ClC-Ka in conjunction with biochemical data. The cytoplasmic domain of the channel ClC-Ka is a homodimer of two structurally similar CBS motif containing proteins. Each subunit exhibits an organization that is common to proteins with a similar fold. Unlike its counterpart in the transporter ClC-5, the ClC-Ka domain does not interact

with adenosine nucleotides. As the equivalent structure of the channel ClC-0 has no affinity for nucleotides either, it remains an important question whether there might be yet undiscovered ligands that bind to the cytoplasmic domains of certain ClC family members and thereby influence the functional behavior of these transport proteins. As ClC-K channels were found to interact with their auxiliary β subunit Barttin, which itself is a putative transmembrane protein, another unresolved question concerns a possible interaction between the large cytoplasmic component of Barttin and the cytoplasmic domain of ClC-Ka (Estevez et al., 2001; Scholl et al., 2006). As the functional behavior of ClC-K channels are to date only poorly characterized, it is not a surprise that we currently lack detailed information on the functional role of their cytoplasmic components.

Apart from its high-resolution structure, the main important feature discovered in this study concerns the oligomeric organization of the ClC-Ka domain, which is likely general for other members of the family. The structure reveals a conserved interaction interface that relates two protein chains by a two-fold symmetric arrangement. The intermolecular interactions are predominantly formed between residues of the CBS2 motifs and bury more than 1500 $Å^2$ of the combined molecular surface. When including coordinated water molecules, most of which are found in a narrow tunnel crossing the interface, the contact surface increases to 2000 $Å^2$. Several observations suggest that the intermolecular interactions are preserved among ClC proteins, as follows. (1) The dimeric relationship in the ClC-Ka domain is very similar to an assembly found in the previously determined ClC-5 domain structure (Meyer et al., 2007). The 2-fold related CBS2 subdomains of the respective dimers superimpose with a root-mean-square deviation (rmsd) of only 1.4 Å, and several residues contributing to the interface are conserved in sequence and structure (Figures 1A and 5A). (2) Biochemical experiments from this study show strong evidence that the oligomeric structure of the ClC-Ka domain represents its structure in solution, and similar experiments in ClC-0 strongly suggest that a pair of cytoplasmic domains of this distantly related channel share the same interaction interface. (3) Crosslinking experiments on full-length ClC-0 in its membrane environment indicate that the observed assembly is also preserved in the context of the transmembrane channel.

The dimeric structures of the cytoplasmic ClC domains are unusual and do not resemble common interaction modes found in other proteins that share a similar fold (Miller et al., 2004; Rudolph et al., 2007; Townley and Shapiro, 2007). These head-to-head or head-to-tail dimers, which we summarize as "conventional interfaces," usually interact via a flat surface formed by the two α helices of each CBS subdomain. Conventional interfaces usually bury a larger surface between 3300 and 4300 $Å^2$ and confer a disk-like appearance to the protein assembly (Figure 5B). Two predominant features distinguish cytoplasmic ClC domains from other CBS motif containing proteins, as follows. (1) The chemical features of the interface defined in the ClC family are not preserved in proteins that oligomerize as "conventional dimers" and vice versa. The respective interfaces thus exhibit a greater degree of sequence conservation and larger hydrophobic patches. (2) The helices of proteins forming conventional dimers are positioned to allow a snug fit upon dimerization, resulting in extensive interactions between both subdomains. Although the individual CBS motifs are generally similar in structure, a difference in their arrangement with respect to each other prevents the formation of an equally large interface for the respective ClC domains. For example, in the case of the cytoplasmic domain of the channel ClC-Ka, this conformational difference is most pronounced (Figure 5B). A partial structure including both CBS motifs superimposes with an rmsd of more than 4.6 Å with the respective regions of the conventional dimer, forming protein TM0935 from *Thermotoga maritima* (whereas the individual CBS motifs of the two proteins would superimpose with an rmsd of only 1.8 Å and 1.2 Å, respectively) (Miller et al., 2004). A similar oligomerization of the ClC-Ka domain would therefore only bury about 1800 $Å^2$ of the common molecular surface as compared to more than 4300 $Å^2$ in the case of TM0935.

The considerably smaller interaction interface found in the structures of the cytoplasmic domains of ClC proteins might reflect the necessity to undergo conformational changes upon transport regulation (Pusch et al., 1997). Such a conformational change has recently been observed in the channel ClC-0. Using Forster resonance energy transfer spectroscopy, the authors of the study described a distance increase between the two C termini in the dimeric channel of more than 20 Å upon common gate closure, a regulatory mechanism which affects both subunits of the homodimeric protein (Bykova et al., 2006; Miller, 1982). Although the structure of the ClC-0 domain was previously determined, its oligomeric organization was not revealed by the crystal structure (Meyer and Dutzler, 2006). When generating a dimeric assembly of the ClC-0 domain using the ClC-Ka domain structure as a template, the resulting dimer shows an interesting feature: despite an overall fit, a dimeric organization as seen in ClC-Ka or ClC-5 appears to be prevented by the local conformation of the protein at the beginning of strand β1. This region contains a sequence stretch that is conserved in ClC-5 and most other ClC proteins (although not in ClC-Ka). The conserved SPF motif (involving the residues Ser 722, Pro 723, and Phe 724 in ClC-0) is well defined in the equivalent region of the transporter ClC-5 but differs in its conformation with respect to ClC-0 (Figure 5C). We have shown that a mutation of Phe 724 destabilizes dimerization in the ClC-0 domain, thereby underlining its role in the interaction interface (Figures 4C and 4D). It thus remains to be shown whether this conformational difference in ClC-0 could represent an alternative conformation that would lead to a change in the dimeric assembly as seen during common gating.

In this study, we have described the dimeric structure of the cytoplasmic domain of the human channel ClC-Ka,

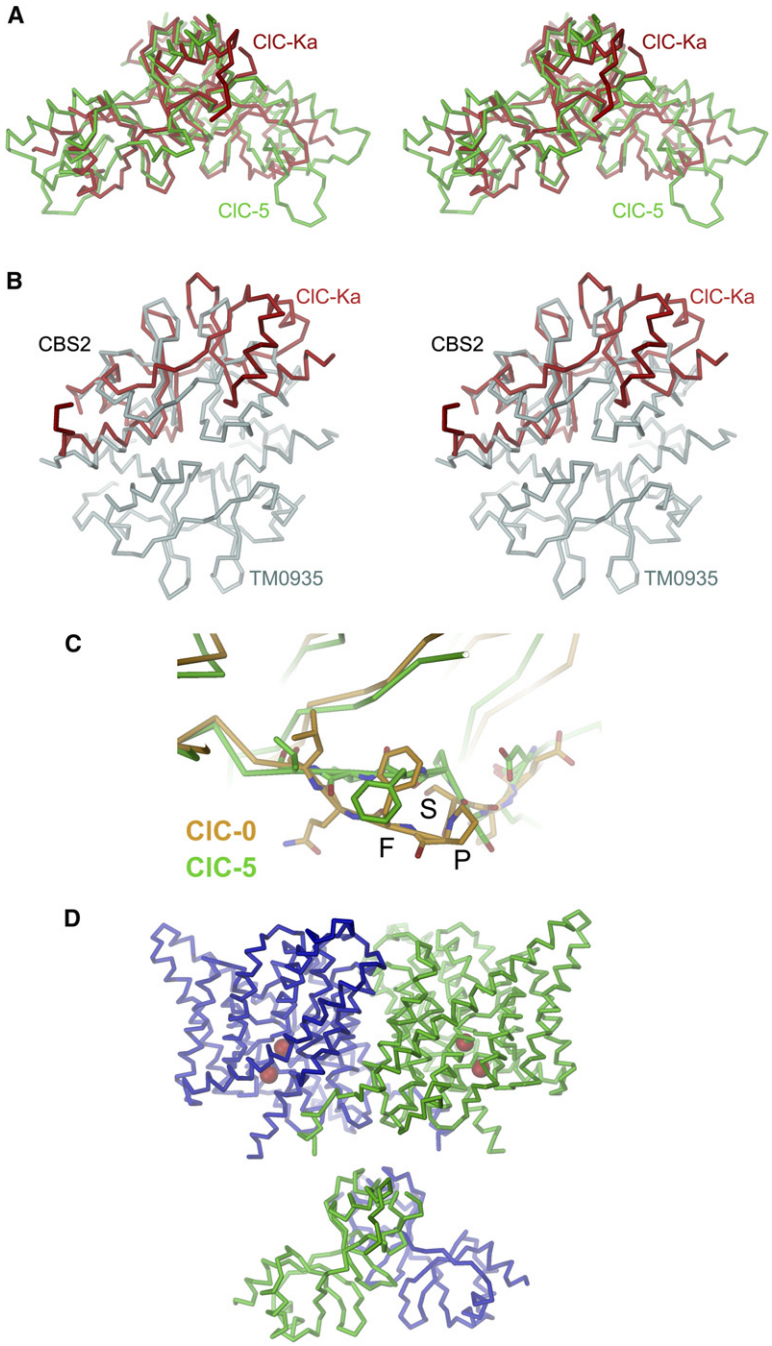



**Figure 5. Structural Comparison of ClC Cytoplasmic Domains**

(A) Stereo view of a superposition of the ClC-Ka domain dimer (red) and the ClC-5 domain dimer (green). The superposition was generated by a least-square fit of the two CBS2 subdomains of the respective dimer. The structures are shown as Cα traces.

(B) Stereo view of a Cα trace of the head-to-tail dimer of the protein TM0935 from *T. maritima* (cyan). The structure of a Cα trace of a single subunit of the cytoplasmic domain from ClC-Ka is shown superimposed (red). The superposition was generated by a least-square fit of the respective CBS2 subdomain of ClC-Ka on the CBS2 subdomain of a single subunit of TM0935.

(C) Superposition of the cytoplasmic domains from ClC-0 (orange) and ClC-5 (green). A section of a single subunit around β strand 1 of CBS2 close to the dimer interface is shown. The view is approximately along the dimer axis. Selected residues including the conserved SPF motif (Ser 722, Pro 723, and Phe 724 in ClC-0) are shown as sticks.

(D) Possible relationship in the full-length ClC protein. Cα trace of the EcClC dimer and the dimeric structure of the cytoplasmic domain of ClC-Ka are shown. The subunits are colored in blue and green, respectively. Bound ions in the EcClC structure are shown as red spheres. The exact relationship between the two components is currently unknown.

and have shown evidence that the observed intermolecular interaction is conserved within members of the ClC family. It is very likely that this quaternary arrangement is preserved in the context of the full-length proteins (Figures 4D and 5D). The next challenging step is to relate these cytoplasmic components to the transmembrane pore and reveal how changes in the domain conformation are transmitted to influence the catalytic unit. With the current study, we provide a structural framework for future investigations which may shed light on these poorly understood mechanisms.

## EXPERIMENTAL PROCEDURES

### Protein Preparation

For expression in *Escherichia coli*, the cytoplasmic domain of the channel ClC-Ka from *Homo sapiens* (GenBank accession number P51800) encompassing residues 533–687 was cloned into the pET28-b+ vector (Novagen) with a C-terminal recognition site for Prescission protease (GE Healthcare) and a hexahistidine tag. *E. coli* BL21 (DE3) cells transformed with the expression construct were grown at 37°C in Luria broth medium containing 50 mg/l kanamycin to an $OD_{600}$ of 1.5. Expression was induced by addition of 0.5 mM isopropyl-D-thiogalactopyranoside (IPTG) at 20°C overnight. Cells

were harvested and lysed by sonication in 50 mM Tris-HCl (pH 7.5), 200 mM NaCl, 5 mM β-mercaptoethanol, 1 mg/ml lysozyme, 20 μg/ml DNase, 1 μg/ml leupeptin, 1 μg/ml pepstatin, and 1 mM phenylmethyl sulfonyl fluoride (PMSF). The lysate was cleared by centrifugation and the protein was purified by affinity chromatography on a nickel chelating Sepharose column (GE Healthcare). The pure protein was cleaved with Prescission protease (GE Healthcare) to remove the hexahistidine tag and dialyzed into TBS buffer (10 mM Tris-HCl [pH 7.5], 200 mM NaCl, 5 mM β-mercaptoethanol). After concentration, the protein was subjected to gel filtration on a Superdex 200 column (GE Healthcare). For preparation of selenomethionine-labeled protein, the bacterial culture was grown in minimal medium containing 50 mg/l Se methionine. The cytoplasmic domain of ClC-0 from *T. marmorata* was cloned into the pET28-b+ vector (Novagen) with a C-terminal recognition site for Prescission protease (GE Healthcare) and a hexahistidine tag and was expressed and purified as previously described (Meyer and Dutzler, 2006). Mutations were inserted with the QuikChange method (Stratagene) and confirmed by sequencing. Mutant proteins were purified following the same protocol as for WT. For crosslinking studies, the respective disulfide bridges formed spontaneously upon exposure to air and were reduced by addition of 250 mM β-mercaptoethanol. The proteins were analyzed by SDS-PAGE, and molecular masses were confirmed by electrospray ionization (ESI) mass spectrometry.

### Crystallization and Crystal Preparation

Crystals of the ClC-Ka cytoplasmic domain were obtained at 4°C by the sitting drop vapor diffusion method. Purified protein with a concentration of 5–10 mg/ml was mixed in a 1:1 ratio with reservoir solution containing 100 mM $MgCl_2$, 50 mM Tris-HCl (pH 8.5), and 20% PEG 4000 and was equilibrated against the reservoir. The reservoir solution of crystals used for the high-resolution data sets and for cocrystallization with ATP contained 100 mM KI instead of $MgCl_2$. Crystals grew after 2 days in space group $P2_1$ (a = 35.3 Å, b = 83.4 Å, c = 58.9 Å, $\alpha = 90°$, $\beta = 92.5°$, $\gamma = 90°$) with two copies of the ClC-Ka domain in the asymmetric unit. Heavy-metal derivatives were obtained by soaking native crystals for 6 hr in mother liquor containing 1 mM methyl mercury chloride, 1 mM Hg(II) chloride, or 300 mM KI. For cryoprotection, the crystals were successively transferred into mother liquor containing increasing concentrations of ethylene glycol to a final concentration of 25%, flash-frozen in liquid propane, and stored in liquid nitrogen.

### Crystallography

Data were collected on a Mar225 CCD detector (Marresearch) at the X06SA beamline at the Swiss Light Source of the Paul Scherrer Institute and on a Mar345 imaging plate on an in-house rotating anode (Nonius) (Table 1). The data were indexed and integrated with DENZO and SCALEPACK and further processed with CCP4 programs (CCP4, 1994; Otwinowski and Minor, 1997). The structure of the human ClC-Ka domain was determined by the multiple isomorphous replacement method making use of three heavy-atom derivatives. Initial low-resolution phases were determined by the single anomalous dispersion method from a data set from a crystal of Se methionylated protein collected at the anomalous absorption edge of Se. The four Se sites in the asymmetric unit were identified with the program SHELXC and -D (Schneider and Sheldrick, 2002). The sites were refined and low-resolution phases were calculated in SHARP (de La Fortelle and Bricogne, 1997). These low-resolution phases allowed the identification of heavy-atom positions in three derivatives by difference-Fourier techniques. The heavy-atom positions were refined in SHARP, and phases were initially improved in Solomon and subsequently by twofold noncrystallographic symmetry (NCS) averaging in DM (Abrahams and Leslie, 1996; Cowtan, 1994; de La Fortelle and Bricogne, 1997). An initial model was built into electron density of a data set at 2.6 Å in O and refined in CNS using strict NCS constraints (Brunger et al., 1998; Jones et al., 1991). The structure was subsequently refined in a data set collected to 1.6 Å using CNS. Initial NCS constraints were subsequently loosened, and restrained individual B factors were refined using CNS (Brunger et al., 1998). The same set of reflections (10%) was set aside for calculation of $R_{free}$ throughout the refinement. The final model contains 2283 protein atoms, 286 water molecules, one $I^-$ ion, and three $Cl^-$ ions with good geometry. All four ions bind at the periphery of the protein and probably do not have any functional significance. This model was used for refinement in a data set at 1.8 Å from a crystal grown in the presence of 5 mM ATP. The structure of the protein in these crystals is virtually identical to the structure in the absence of ATP, and no residual density for bound ATP could be identified (Table 1). The interaction interfaces between subunits were calculated in CNS (Brunger et al., 1998).

### Analytical Ultracentrifugation

Sedimentation velocity experiments were performed with a Beckman model XL-I analytical ultracentrifuge with an An 50-Ti rotor at 4°C. Double-sector centerpieces were filled with 400 μl of protein samples and 420 μl of TBS buffer. Data were acquired at 280 nm in continuous-scan mode in 0.003 cm intervals at a rotor speed of 40,000 rpm. Sedimentation analysis of the ClC-Ka cytoplasmic domain and its mutants F636D and A597D was performed at protein concentrations of 15 μM. For the ClC-0 cytoplasmic domain, WT protein and the domain mutants P724R, V727R, K567A, and A766R were used at concentrations of 17.5, 35, 70, and 140 μM. Data analysis was performed with the c(S) module of Sedfit (Schuck et al., 2002). The buffer parameters, partial-specific volume of the protein, and the corrected sedimentation coefficients $S^\circ_{20W}$ were calculated by using Sednterp (Laue et al., 1992). The $S^\circ_{20W}$ values for the WT ClC-Ka domain (2.34 S) and for the mutant F636D (1.81 S) are consistent with a dimeric and a monomeric molecular assembly, respectively. Also in the case of the ClC-0 domain, the $S^\circ_{20W}$ values for the WT ClC-0 domain at 140 μM protein concentration (4.26 S) and for the mutant V727R at 17.5 μM protein concentration (2.46 S) match the expected values for a dimer and a monomer, respectively. For data used in this analysis, the root-mean-square deviations of all the fits were between 0.005 and 0.01.

### Expression of ClC-0 in *X. laevis* Oocytes and Western Blot Analysis

ClC-0 from *T. marmorata* was cloned into the pTLN vector as a C-terminal hexahistidine fusion protein for expression in *X. laevis*. Mutation L741C was introduced with the QuikChange method (Stratagene) and confirmed by DNA sequencing. After the linearization of plasmid DNA by MluI, capped cRNA was transcribed with the mMessage mMachine kit (Ambion) and purified with the RNeasy kit (QIAGEN). For expression, 20 ng of RNA was injected into the defolliculated oocytes. The oocytes were kept at 17°C in Barth's solution for 3 days after injection. Oocytes were homogenized in lysis buffer containing 5 mM Tris-HCl (pH 6.9), 250 mM sucrose, 0.5 mM EDTA, 1 μg/ml leupeptin, 1 μg/ml pepstatin, and 1 mM PMSF. The lysate was centrifuged two times at 100 × g at 4°C to remove large cellular debris. For oxidative crosslinking, the supernatant containing the membrane fraction was treated with 100 mM sodium tetrathionate for 30 min at 4°C and then separated on an 8% SDS-polyacrylamide gel. Western blot analysis with anti-$His_6$-peroxidase monoclonal antibody was performed according to the manufacturer's instructions (Roche Applied Science). Proteins were detected with Immobilon western chemiluminescent HRP substrate (Millipore) and BioMax XAR photographic film (Kodak).

## ACKNOWLEDGMENTS

We are grateful to the staff of the X06SA beamline for their support during data collection, S. Chesnov and P. Hunziker for help with mass spectrometry, E. Hänsenberger for preparation of the *Xenopus* oocytes, B. Blattmann for help with crystal screening, Sebastian Meyer for scientific advice and for comments on the manuscript, members of the Dutzler lab for help in all stages of the project, and Thomas Jentsch for providing the ClC-Ka and ClC-0 c-DNA. Data collection was

performed at the Swiss Light Source of the Paul Scherrer Institute. This work was supported by a grant from the Swiss National Science Foundation and the National Center of Competence in Research, Structural Biology program. S.M. is affiliated with the Molecular Life Sciences Ph.D. program of the University/ETH Zürich. The authors declare that they have no competing financial interest.

Received: March 4, 2007
Revised: April 4, 2007
Accepted: April 5, 2007
Published: June 12, 2007

## REFERENCES

Abrahams, J.P., and Leslie, A.G. (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta Crystallogr. D Biol. Crystallogr. *52*, 30–42.

Bateman, A. (1997). The structure of a domain common to archaebacteria and the homocystinuria disease protein. Trends Biochem. Sci. *22*, 12–13.

Bennetts, B., Rychkov, G.Y., Ng, H.L., Morton, C.J., Stapleton, D., Parker, M.W., and Cromer, B.A. (2005). Cytoplasmic ATP-sensing domains regulate gating of skeletal muscle ClC-1 chloride channels. J. Biol. Chem. *280*, 32452–32458.

Biemans-Oldehinkel, E., Mahmood, N.A., and Poolman, B. (2006). A sensor for intracellular ionic strength. Proc. Natl. Acad. Sci. USA *103*, 10624–10629.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR System: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Bykova, E.A., Zhang, X.D., Chen, T.Y., and Zheng, J. (2006). Large movement in the C terminus of CLC-0 chloride channel during slow gating. Nat. Struct. Mol. Biol. *13*, 1115–1119.

CCP4 (Collaborative Computational Project, Number 4) (1994). The CCP4 suite: programs for X-ray crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Cowtan, K. (1994). An automated procedure for phase improvement by density modification. Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography *31*, 34–38.

de La Fortelle, E., and Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. In Methods in Enzymology, C.W. Carter and R.M. Sweet, eds. (New York: Academic Press), pp. 492–494.

Dutzler, R. (2006). The ClC family of chloride channels and transporters. Curr. Opin. Struct. Biol. *16*, 439–446.

Dutzler, R., Campbell, E.B., Cadene, M., Chait, B.T., and MacKinnon, R. (2002). X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. Nature *415*, 287–294.

Estevez, R., Boettger, T., Stein, V., Birkenhager, R., Otto, E., Hildebrandt, F., and Jentsch, T.J. (2001). Barttin is a $Cl^-$ channel β-subunit crucial for renal $Cl^-$ reabsorption and inner ear $K^+$ secretion. Nature *414*, 558–561.

Ignoul, S., and Eggermont, J. (2005). CBS domains: structure, function, and pathology in human proteins. Am. J. Physiol. Cell Physiol. *289*, C1369–C1378.

Jentsch, T.J., Stein, V., Weinreich, F., and Zdebik, A.A. (2002). Molecular structure and physiological function of chloride channels. Physiol. Rev. *82*, 503–568.

Jentsch, T.J., Neagoe, I., and Scheel, O. (2005). CLC chloride channels and transporters. Curr. Opin. Neurobiol. *15*, 319–325.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Laue, T., Shah, B., Ridgeway, T., and Pelletier, S. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. In Analytical Ultracentrifugation in Biochemistry and Polymer Science, S. Harding, A. Rowe, and J. Horton, eds. (Cambridge, UK: The Royal Society of Chemistry), pp. 90–125.

Maduke, M., Williams, C., and Miller, C. (1998). Formation of CLC-0 chloride channels from separated transmembrane and cytoplasmic domains. Biochemistry *37*, 1315–1321.

Meyer, S., and Dutzler, R. (2006). Crystal structure of the cytoplasmic domain of the chloride channel ClC-0. Structure *14*, 299–307.

Meyer, S., Savaresi, S., Forster, I.C., and Dutzler, R. (2007). Nucleotide recognition by the cytoplasmic domain of the human chloride transporter ClC-5. Nat. Struct. Mol. Biol. *14*, 60–67.

Miller, C. (1982). Open-state substructure of single chloride channels from *Torpedo electroplax*. Philos. Trans. R. Soc. Lond. B Biol. Sci. *299*, 401–411.

Miller, M.D., Schwarzenbacher, R., von Delft, F., Abdubek, P., Ambing, E., Biorac, T., Brinen, L.S., Canaves, J.M., Cambell, J., Chiu, H.J., et al. (2004). Crystal structure of a tandem cystathionine-β-synthase (CBS) domain protein (TM0935) from *Thermotoga maritima* at 1.87 Å resolution. Proteins *57*, 213–217.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *267*, 307–326.

Pusch, M., Ludewig, U., and Jentsch, T.J. (1997). Temperature dependence of fast and slow gating relaxations of ClC-0 chloride channels. J. Gen. Physiol. *109*, 105–116.

Rudolph, M.J., Amodeo, G.A., Iram, S.H., Hong, S.P., Pirino, G., Carlson, M., and Tong, L. (2007). Structure of the Bateman2 domain of yeast Snf4: dimeric association and relevance for AMP binding. Structure *15*, 65–74.

Sanner, M.F., Olson, A.J., and Spehner, J.C. (1996). Reduced surface: an efficient way to compute molecular surfaces. Biopolymers *38*, 305–320.

Schneider, T.R., and Sheldrick, G.M. (2002). Substructure solution with SHELXD. Acta Crystallogr. D Biol. Crystallogr. *58*, 1772–1779.

Scholl, U., Hebeisen, S., Janssen, A.G., Muller-Newen, G., Alekov, A., and Fahlke, C. (2006). Barttin modulates trafficking and function of ClC-K channels. Proc. Natl. Acad. Sci. USA *103*, 11411–11416.

Schuck, P., Perugini, M.A., Gonzales, N.R., Howlett, G.J., and Schubert, D. (2002). Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. Biophys. J. *82*, 1096–1111.

Scott, J.W., Hawley, S.A., Green, K.A., Anis, M., Stewart, G., Scullion, G.A., Norman, D.G., and Hardie, D.G. (2004). CBS domains form energy-sensing modules whose binding of adenosine ligands is disrupted by disease mutations. J. Clin. Invest. *113*, 274–284.

Townley, R., and Shapiro, L. (2007). Crystal structures of the adenylate sensor from fission yeast AMP-activated protein kinase. Science *315*, 1726–1729.

Vanoye, C.G., and George, A.L., Jr. (2002). Functional characterization of recombinant human ClC-4 chloride channels in cultured mammalian cells. J. Physiol. *539*, 373–383.

### Accession Numbers

The ClC-Ka domain structure has been deposited in the Protein Data Bank under ID code 2PFI.

# 3 General Discussion and Outlook

The cytoplasmic domains in eukaryotic ClC proteins have been recognized to be intimately involved in regulatory processes. One of these processes is the regulation of ClCs in response to nucleotide binding that has been shown to have opposite effects on ion transport in different ClC proteins. While in the muscle channel ClC-1 intracellular ATP has an inhibitory effect (Tseng, Bennetts et al. 2007), in the human transporter ClC-5 ATP increases the probabilty for an active, open state (Zifarelli and Pusch 2009). Although the chemistry of nucleotide binding has been revealed from the crystal structure of the ClC-5 domain (Meyer, Savaresi et al. 2007), structural differences i.e. conformational changes expected to happen in the intracellular domains upon ligand binding are still not known. A probably related mechanism of regulation is the process of common gating in ClC channels that has been recently associated with the movements in the cytoplasmic domains (Bykova, Zhang et al. 2006). It has been demonstrated that the closure of the common gate in ClC-0 coincides with the separation of the cytoplasmic domains by a distance up to 20 Å. In order to fully understand the described regulatory phenomena a structural basis for the organization of cytoplasmic domains in the context of full length ClC proteins is ultimately required.

It has been anticipated that the cytoplasmic domains in ClC proteins extend the two-fold symmetry of the transmembrane domains and that the interaction between them is important for the structure and function of these dimeric proteins. However, the previous structural and biochemical studies on the cytoplasmic domains of ClC-0 and ClC-5 have left their mode of dimerization ambiguous (Meyer and Dutzler 2006; Meyer, Savaresi et al. 2007). In this work the dimeric organization of the cytoplasmic domains in eukaryotic ClC proteins has been clarified through the examination of the domain of the human channel ClC-Ka (Markovic and Dutzler 2007). The structural and biochemical analyses of the ClC-Ka domain confirmed a novel mode of dimerization not found in other CBS domain containing proteins. This study has also suggested that the dimeric interaction is conserved in the eukaryotic family members and likely preserved in the context of the full-length proteins. Therefore, the structure of the ClC-Ka domain dimer serves as an initial structural basis for comprehending possible domain movements asociated with the common gating in ClC-0 and it allows the design of future functional experiments.

The interaction in the ClC-Ka domain dimer is formed mainly through the CBS2 subdomains and leads to a V-shaped dimer with the combined surface of the interface of

around 1500 Å$^2$. This considerably smaller interface and a different arrangement of subunits than found in other CBS domain dimers might indicate the necessity to undergo conformational changes during the regulation of ion transport. Although mainly composed of hydrophobic residues, the interface contains a number of water molecules which build a tunnel in the lower part of the contact area. This finding could suggest a dissociation of the dimer in this part of the interface during the domain movement. Other interesting features of the ClC cytoplasmic domains discovered in this work are different affinities of the interaction found for the ClC-Ka and the ClC-0 domain by sedimentation analysis. While the ClC-Ka domain showed dimeric behavior independent of the tested protein concentrations, the ClC-0 domain dissociates at low protein concentration indicating lower affinity. This would mean that while the general dimeric arrangement of the cytoplasmic domains is conserved within eukaryotic ClC proteins, it might be energetically tuned in each specific protein depending on the regulatory function it needs to fulfill. The weaker interaction in the ClC-0 dimer agrees with the movements happening during the common gate closure while the stable dimer of the ClC-Ka cytoplasmic domain may be important for the unique function of this channel. ClC-K channels are the only ClCs that contain a valine residue instead of the gating glutamate, which implies that they do not contain a protopore gate that closes the individual pores. In case the tight interaction between the cytoplasmic domains also stabilizes the open state of the common gate (which has not yet been confirmed to exist in ClC-K channels), the channels may be constitutively open. Such non-gated, ohmic “leak” channels are suited to support the transepithelial transport over a large range of membrane potentials, which is the function of ClC-K channels in the epithelia of kidney and inner ear. The fact that ClC-K channels have a regulatory subunit Barttin, that has been implicated to modulate gating and conduction in heterologous expression systems (Scholl, Hebeisen et al. 2006), could also mean that ion transport in the ClC-K channels is not regulated by common mechanisms found in other ClC channels, but instead by the interaction with the accessory subunit. The ultimate goal in explaining the molecular basis for the regulatory role of the β-subunit Barttin is structure determination of the ClC-K channel – Barttin complex.

Although the structure of the ClC-Ka domain has made an important advance towards revealing structural relationship between the cytoplasmic domains in eukaryotic ClCs, the question of how the domains would be positioned with respect to the catalytic membrane domains and the pore remains unclear. In the hypothetical arrangement of the full-length protein shown in Figure 3.1, the CBS2 subdomains of the dimer are positioned in proximity to the membrane domain. Since the N-terminus wraps around CBS2, the connection with the last helix of the transmembrane part could be easily established. This hypothesis is supported by split channel experiments that showed that functional ClC-0 and ClC-1 channels could be reconstituted upon co-expression of the part containing the

transmembrane domain and the CBS1 subdomain and the part comprising CBS2 (Schmidt-Rose and Jentsch 1997; Maduke, Williams et al. 1998). These results have suggested that CBS2 is closer to the membrane and thus necessary to form the interaction between the cytoplasmic and the transmembrane domain.

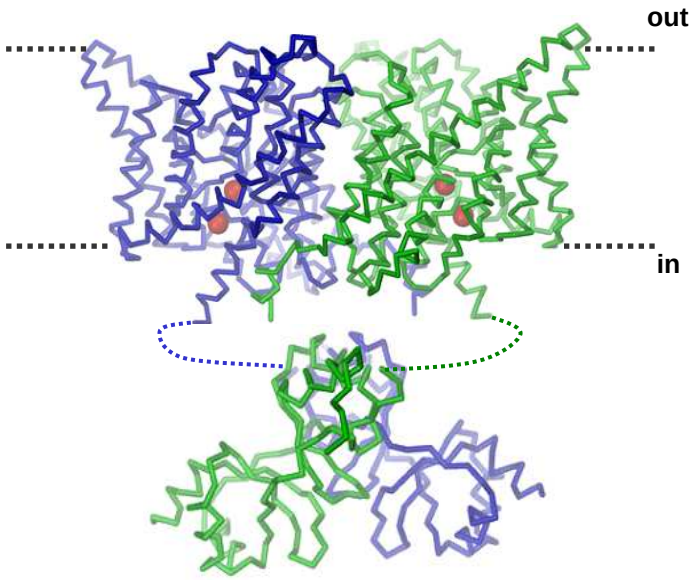



**Figure 3.1. Possible relationship in the full-length eukaryotic ClC protein.**
Cα trace of the EcClC dimer and the dimeric structure of the cytoplasmic domain of ClC-Ka are shown. The subunits are colored in blue and green, respectively. Bound ions in the EcClC structure are shown as red spheres. Dotted lines present the linker between the last transmembrane helix and the N-terminus of the cytoplasmic domain.

To resolve the organization and arrangement of the catalytic and intracellular domains it is crucial to determine the structure of ClC proteins that contain both components. During the course of this PhD project an effort has been taken in establishing *Pichia pastoris* expression system for eukaryotic ClC homologs (see Appendix 4.3). However, the performed experiments showed that the expression of eukaryotic ClC proteins still needs to overcome experimental challenges to be useful for structural studies. In the future, besides the *P. pastoris* other expression systems like baculovirus-mediated insect cell expression and expression in mammalian cells will have to be explored. To increase the chance of finding eukaryotic candidates suitable for crystallization and structure determination, the expression screening should be performed in a high-throughput manner including testing of different proteins, tags and modifications of individual constructs.

In several cases it has proven useful to turn to prokaryotic homologs for structural studies of membrane proteins (Doyle, Morais Cabral et al. 1998; Luecke, Schobert et al. 2001; Locher, Lee et al. 2002; Yamashita, Singh et al. 2005; Hilf and Dutzler 2008). These

proteins are generally more readily expressed, biochemically more stable and more amenable to crystallization. There are a number of prokaryotic ClC homologs that share the molecular architecture of eukaryotic proteins, containing both the catalytic domain with all the conserved features of the selectivity filter and the cytoplasmic domains with both CBS subdomains. In this work an extensive screening of prokaryotic ClC homologs for expression in *E. coli* has been carried out. The critical factor in the screening was the expression of the membrane proteins under the control of a tunable pBAD promoter and the use of C-terminally fused GFP, which serves as a folding indicator and allows a straightforward quantification prior to purification. In this broad screen of conditions encompassing 27 different homologs only one protein proved to be suitable for crystallization experiments.

The selected ClC homolog, Rm1ClC, originates from the bacterium *Ralstonia metallidurans*, an extremophile capable of growing in high concentrations of heavy metals. The genome of this prokaryote provides a rich source of different membrane transport proteins (von Rozycki, Nies et al. 2005). The sequence of Rm1ClC contains the conserved features of the ClC family and it is, additionally, unusually compact due to the truncation of loops connecting the helices of the transmembrane domain and the short CBS domain. The Rm1ClC protein is highly expressed and stable in mild detergents suitable for crystallization. However, despite having all the characteristics that makes it a promising target for crystallization, full-length Rm1ClC has not given crystals during the course of this PhD project despite the exhaustive screening of conditions. It might be interesting to speculate about the underlying cause for this behavior. One possibility is that the protein exists in multiple energetically equivalent conformations which would prohibit crystal lattice formation. While the transmembrane domain and the cytoplasmic domains as separate entities are expected to be relatively rigid, the above discussed regulatory mechanisms in eukaryotic ClCs would imply that the regions linking the two domains might confer flexibility. This hypothesis is to some extent supported by the successful crystallization of the transmembrane and the cytoplasmic domain of Rm1ClC in isolation. The fact that both the transmembrane and the cytoplasmic domain are stable when expressed independently suggests that the interaction between these domains in the context of the full-length protein might be weak, thus promoting flexibility.

The transmembrane domain of Rm1ClC crystallized readily, but at the same time revealed the challenges often observed in membrane protein crystallography. The obtained crystals diffracted X-rays only to low resolution and were overall resistant to improvement by variations of the detergents used for crystallization and modifications of the expression construct. When aiming for the elucidation of novel properties and minor structural differences between a new ClC family member and the currently known proteins, high resolution structure is essential. Therefore, future crystallization experiments on the

Rm1ClC transmembrane domain will include co-crystallization with the selected binders (see Appendix 4.1), which might contribute to a better crystal packing and provide crystals diffracting to higher resolution.

The structure of the Rm1ClC cytoplasmic domain determined in this work is the second structure of a CBS domain originating from a prokaryotic ClC protein. While both prokaryotic structures, the Rm1ClC and the PtClC cytoplasmic domain (Looser 2008) have a typical CBS fold, they do not show dimeric arrangement found to be conserved in the eukaryotic ClC proteins. Both prokaryotic domains are monomers in solution and the analyzed contacts in the crystals of both proteins do not suggest likely dimeric organizations in the context of the full-length proteins. This finding directly opens a question on the role of these intracellular components in prokaryotic ClC proteins. It will have to be shown whether they play a solely structural role in keeping the R helix in a certain conformation required or whether they interact in order to regulate ion transport in the protein. An answer to this question will require structural information on the full-length protein. In order to reach this long standing goal, future experiments will have to include co-crystallization experiments of the Rm1ClC full-length protein with selected binders, assuming that some of them might stabilize the cytoplasmic domains in a conformation suitable for crystallization. Another way to reveal the arrangement of the cytoplasmic domains within the prokaryotic ClC proteins could come from electron paramagnetic resonance (EPR) spectroscopy. EPR spectroscopy allows indirect determination of structural organization of proteins by measuring distances between introduced spin labels. Labeling of the cytoplasmic domains of Rm1ClC might provide information on their structural relationship in the full-length protein, but also on their flexibility. Other indirect methods, like chemical cross-linking and FRET are valuable alternatives for the analysis of the cytoplasmic domain assembly.

One of the important goals will be to clarify whether the cytoplasmic domains of prokaryotic ClCs bind ligands, like some of their eukaryotic counterparts. The binding of nucleotides seems unlikely in light of the known structure of the Rm1ClC and PtClC domain. However, some other molecules or ions might serve as ligands for these proteins. The CBS domains that bind ions and act as ion sensors are found in the transporters MgtE (Hattori, Iwase et al. 2009) and OpuA (Biemans-Oldehinkel, Mahmood et al. 2006). In the crystal structure of the Rm1ClC CBS domain chloride and zinc ions are detected at prominent positions. The affinity of ion binding will have to be tested by titration calorimetry and biochemical assays. The screening for other potential ligands will have to be continued, which might ultimately lead to the discovery of molecules that induce conformational changes in the domain and regulate ion transport. At this point, structural and functional analyses will have to go hand in hand to explain the mechanism of regulation.

A future focus will also have to be on a functional characterization of Rm1ClC by electrophysiological recordings and transport assays. It will be important to determine the affinities for chloride ions in the selectivity filter. From combined functional experiments it should be possible to draw conclusions about the rate of ion transport and reveal whether this protein functions as a channel or as a transporter. The presence of valine residue at the position of $Glu_{int}$ does not unambiguously allow to deduce whether Rm1ClC functions as channel.

In summary, in the context of this thesis a prokaryotic ClC protein containing CBS cytoplasmic domains has been established as promising model for structural and functional studies. The first structural insights into its components have already been obtained. Future experiments on Rm1ClC are expected to establish the role of cytoplasmic domains in prokaryotic ClC proteins and provide some analogy for understanding the regulatory mechanisms in the eukaryotic ClC family members.

# 4 Appendix

## 4.1 Design of antibodies against Rm1ClC protein

### 4.1.1 Binding molecules that facilitate crystallization

In X-ray structure determination of membrane proteins the main obstacle after providing sufficient amounts of a target protein is to obtain well-ordered three dimensional crystals. The amphipatic nature of membrane proteins often makes their crystallization difficult. Membrane proteins contain large hydrophobic surfaces that upon solubilization are covered by detergent molecules and only relatively small hydrophilic regions needed for establishing protein-protein contacts in a three-dimensional crystal lattice. In addition, membrane transport proteins often exist in multiple conformations, which is another unfavorable property for crystal formation.

A strategy for obtaining well-ordered crystals is the use of crystallization chaperons that bind to the target protein and increase its crystallization probability (Koide 2009). These crystallization chaperons enlarge the hydrophilic part of membrane proteins thereby providing additional surface for crystal contacts and reduce a conformational heterogeneity in the target by binding to a specific conformation. Several types of binders are nowadays used to support crystallization of proteins including antibody fragments, designed ankyrin repeat proteins (DARPINs), affibodies (Hogbom, Eklund et al. 2003) etc.

Antibody derived binders like Fabs, Fvs and single chain Fvs (scFvs) are still the most commonly used chaperones in the crystallization of membrane proteins (Figure 4.1.1A). Fabs are composed of variable and constant parts of the light chain and of the N-terminal halve of the heavy chain of IgGs. They can be generated by partial proteolytic digestion of IgGs with papain. Fvs are composed of the variable parts of heavy and light chains and are usually linked via a short peptide to provide stable scFvs. Antibody fragments can be generated either by a combination of animal immunization and hybridoma technology or by *in vitro* selection from engineered libraries by display techniques. Traditional selection approach by animal immunization is often combined with the cloning and expression of Fab and Fv fragments in *E. coli*.

Over the last decade, many important membrane protein structures were determined in the complex with antibody fragments. For instance, the crystallization of the ClC homolog from *E. coli*, EcClC, in the complex with a Fab fragment has enabled its

structure determination at 2.5 Å (Dutzler, Campbell et al. 2003). The improvement of diffraction quality was due to the increased hydrophilic surface (Figure 4.1.2A).

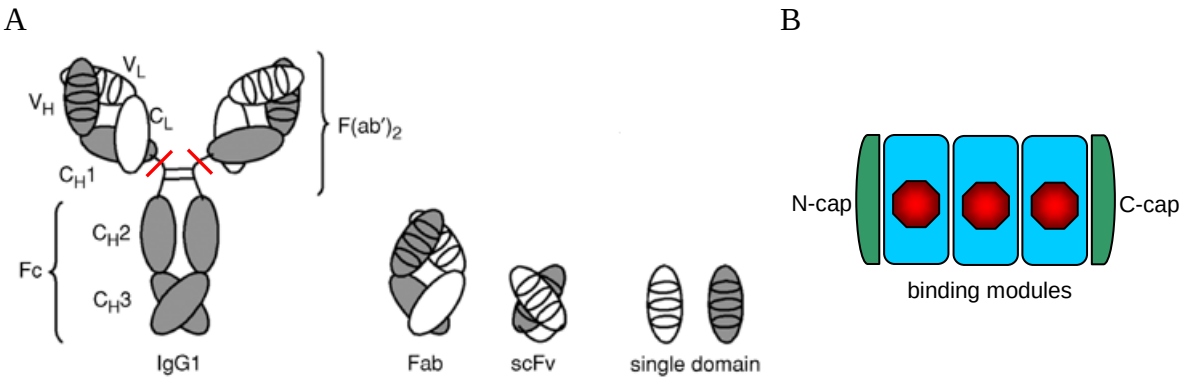



**Figure 4.1.1. Schematic representation of antibody fragments and DARPINs.**
(A) IgGs are part of the immune system and created through hybridoma technology. They consist of two chains, heavy (grey) and light (white) chain. Interaction between the chains and several cysteine bridges stabilize the structure. Both chains contain constant (labeled C) and variable regions (labeled V). Variable regions carry the complementary defining regions (CDRs) that are important for antigen recognition. Fab fragments are generated by proteolytic cleavage in the linker region (red lines).
Figure derived from: http://www.currentprotocols.com/protocol/ps0609
(B) Modular composition of DARPINs. The binding modules (N=3 is shown) have a defined structure encoded by a consensus sequence. Residue variations at restricted positions (red) assemble the library for ribosome display. The N- and C- terminal caps shield the hydrophobic core and generate a stable protein.

DARPINs are synthetic binding proteins composed of ankyrin repeat modules (Figure 4.1.1B) (Binz, Amstutz et al. 2004). They are selected from large libraries by ribosome or phage display. Their advantage compared to antibodies is that they are formed of a single polypeptide chain free of disulfide bonds, have high stability and high expression level in *E. coli*..

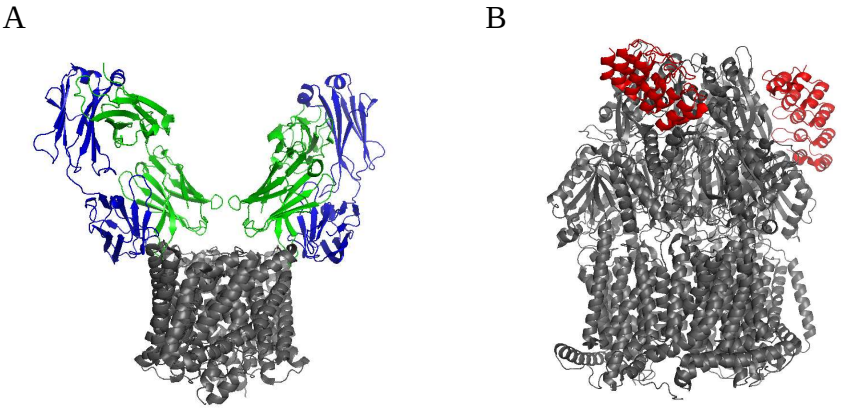



**Figure 4.1.2. Crystallization chaperons in membrane protein structure determination.**
**(A)** Crystallization of complexes of membrane proteins with chaperones can increase the hydrophilic surface and be essential for crystal contacts like in EcClC (grey) crystallized with Fabs (green – heavy, blue – light chain) (1ots). **(B)** Crystallization chaperon can also fix a certain conformation and allow structure determination of distinct transport states like in AcrB (grey) in complex with DARPINs (red) (2j8s).

So far, the use of DARPINs in membrane protein crystallization has been reported only once. Crystallization of the multidrug exporter AcrB with a DARPIN chaperone resulted in a higher resolution of this secondary active transport protein (Sennhauser, Amstutz et al. 2007). The improved crystallization properties were probably due to the selective binding and stabilization of one conformation in the transport cycle of AcrB. The formation of novel crystal contacts by an increase of the polar surface was in this case not observed.

### 4.1.2 Production and identification of positive hybridoma cell lines

The failed crystallization of the Rm1ClC full-length protein and the poorly diffracting crystals of the Rm1ClC transmembrane domain have motivated the production of specific binding proteins for future co-crystallization experiments. The production of monoclonal antibodies was chosen since this approach has so far been most successful in various membrane protein structure determinations.

All experiments were carried out in collaboration with CEPower (Wädenswil, Switzerland), a company specialized for the generation of monoclonal antibodies. Both, the full length and the transmembrane domain of Rm1ClC were used as antigens for the immunization of mice. Three mice were immunized with each antigen (total 6 mice). The immunization was repeated three times before the isolation of the mouse spleen. In order to choose the mouse that shows the best immune response for each antigen, the polyclonal sera from the 6 mice were tested for binding to the antigens by ELISA. Interestingly, it was observed that the immunization with the full length protein gave a stronger immune response than the immunization with the transmembrane domain. After selecting one mouse for each construct, the B-lymphocytes from the whole spleen were fused with myeloma cells to form immortalized hybridoma cells expressing IgGs. The fusion was carried out by standard methods and confirmed by selection on hypoxanthine aminopterin thymidine (HAT) medium. HAT medium is a selection medium where the drug aminopterin inhibits the dihydrofolate reductase, which results in a block of the *de novo* nucleotide synthesis. The synthesis of thymine and guanine can be bypassed by supplying both molecules in the medium. During this selection, only fused cells carrying the immortalization from myeloma cells and the hypoxanthine-guanine phosphoribosyltransferase (HGPRT) from B-lymphocytes can survive.

About 960 stable hybridoma cell lines were isolated for each construct. In order to select the lines secreting monoclonal antibodies against Rm1ClC, the supernatants from all cultures were screened by ELISA for a reaction against the protein. In the first round of selection the ELISA was performed only with the respective construct used for the

immunization i.e. the hybridomas obtained by fusing the spleen of the mouse immunized with the full-length Rm1ClC were tested against the full-length protein. By this procedure 52 clones were selected for antibodies selected against the full length protein, and 20 clones for ones selected against the transmembrane domain. The criterion for selection was an ELISA signal higher than 2 times over the background.

The selected set of positive hybridoma cell lines was in the second round analyzed in more detail. In order to classify the selected monoclonal antibodies based on their binding epitope on Rm1ClC three different constructs were tested by ELISA: full length protein, transmembrane domain and cytoplasmic domain. The results of these experiments are presented in Figure 4.1.1. A strong variation in the ELISA signals coming from different antibodies can be noticed. Although in the first round of selection all selected clones gave signal higher than two times over the background, for many of them this result was not reproduced. The variation of signals recorded in different ELISAs for the same samples has been previously observed (Warmuth 2009). Nevertheless, the number of clones giving reproducibly strong signals was still satisfying.

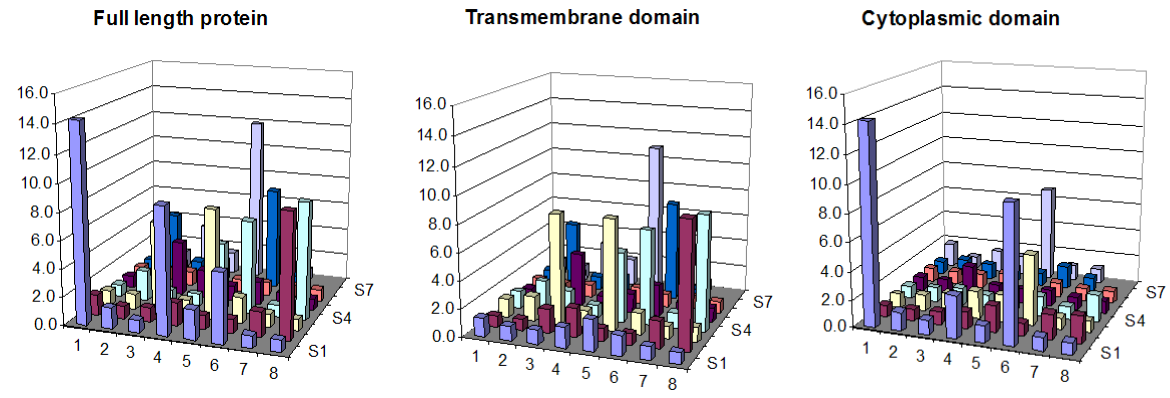



**Figure 4.1.1. Antibody binding against Rm1ClC full-length, transmembrane and cytoplasmic domain.** A set of monoclonal antibodies from 70 hybridoma cell lines was tested for their binding properties by ELISA. Each figure represents an ELISA plate coated with a different antigen. The order of hybridoma supernatant in all three plates is the same. The relative intensity of the chemiluminescence is labeled on the y axis.

It is interesting to analyze the interaction of the monoclonal antibodies with the three different constructs of Rm1ClC. The antibodies binding to the full length protein can be divided into two groups depending on their ability to either bind to the transmembrane or the cytoplasmic domain. Out of 14 clones that gave a strong signal (> 3×background) with the full length protein, ten of them also bind to the transmembrane domain and only four of them bind to the cytoplasmic domain. This initial information on binding regions can be very useful for later co-crystallization trials. The binders to the transmembrane domain are promising candidates for improving the crystal quality since their binding

could enlarge the polar surface of the protein and thereby lead to stronger crystal contacts. The binders to the cytoplasmic domain, in contrast, could be valuable in obtaining crystals of the full length protein by stabilizing the potentially flexible cytoplasmic domains in a specific conformation.

Selected clones that showed a high signal against Rm1ClC were in addition characterized by Western blotting (Figure 4.1.2). The characterization of binding against the denatured protein by Western blot analysis allows the discrimination between sequence epitopes, which show a signal in a Western blot analysis, and structural epitopes, which do not show a signal since the structural epitope is lost on a denaturing gel. It is assumed that antibodies suitable for co-crystallization attempts should form a stable and rigid complex with their native antigen. This implies that they recognize a discontinuous i.e. structural epitope. Sequence epitopes can be unstructured N- and C-terminal tails as well as unstructured loop regions.

Purified Rm1ClC full length protein and the transmembrane domain were separated by SDS-PAGE and blotted on a PVDF membrane. After blocking, the membrane was incubated with the antibody containing supernatants of the selected hybridoma cells, followed by incubation with the secondary antibody and detection of chemiluminescence. Out of 30 tested clones, 13 showed a signal on the Western blot, which suggests that they recognize sequence specific epitopes. It was observed that the antibodies that gave the strongest signals in the ELISA were mostly Western positive.

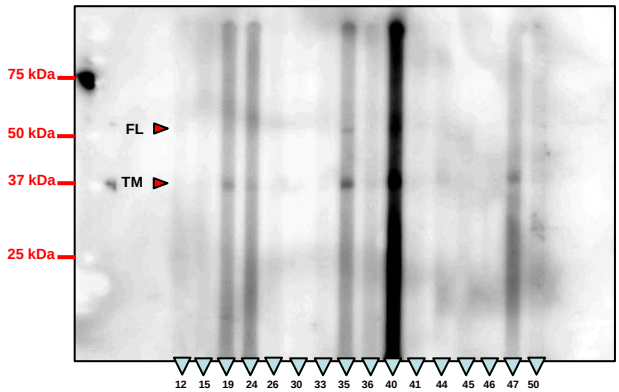



**Figure 4.1.2. Epitope characterization of selected monoclonal antibodies.**
Full-length Rm1ClC (FL) and its transmembrane domain (TM) were separated by SDS-PAGE and blotted on the PVDF membrane. Each lane labeled with the blue triangle was incubated with the indicated antibody of the selected F1 hybridoma cell lines. The Western positive clones i.e. the once that recognize sequence epitopes are: 19, 24, 35, 40, 47 and 50.

After the described analyses, a set of 30 hybridoma lines was selected for further work (Table 4.1). From these hybridoma cell lines recombinant Fab fragments will be generated by cloning of encoding genes. Cloning and expression in *E. coli* should allow

efficient production and purification of recombinant fragments. Purified Fabs will be investigated for binding to Rm1ClC and, if promising will be used for co-crystallization experiments.

| | Clone | ELISA | Western | Binding to |
|---|---|---|---|---|
| 1 | F1-1 | strong | + | FL, Dom |
| 2 | F1-3 | weaker | | FL |
| 3 | F1-12 | weaker | | FL, TM |
| 4 | F1-15 | strong | | FL, TM |
| 5 | F1-18 | weaker | | FL |
| 6 | F1-19 | strong | + | FL, TM |
| 7 | F1-20 | weaker | | FL |
| 8 | F1-22 | weaker | | FL |
| 9 | F1-24 | strong | + | FL, TM |
| 10 | F1-25 | strong | + | FL, Dom |
| 11 | F1-26 | weaker | | FL, TM |
| 12 | F1-28 | weaker | | FL |
| 13 | F1-30 | weaker | | FL, TM |
| 14 | F1-33 | weaker | | FL, TM |
| 15 | F1-35 | strong | + | FL, TM |
| 16 | F1-36 | strong | | FL, TM |
| 17 | F1-40 | strong | + | FL, TM, Dom |
| 18 | F1-41 | strong | | FL, Dom |
| 19 | F1-43 | weaker | | FL, Dom |
| 20 | F1-44 | strong | | FL, TM |
| 21 | F1-45 | weaker | | FL, TM |
| 22 | F1-47 | strong | + | FL, TM |
| 23 | F1-50 | weaker | + | FL, TM |
| 24 | F3-4 | weaker | | TM |
| 25 | F3-8 | strong | + | FL, TM |
| 26 | F3-12 | strong | + | TM |
| 27 | F3-13 | weaker | + | TM |
| 28 | F3-14 | strong | + | FL, TM |
| 29 | F3-15 | strong | + | TM |
| 30 | F3-18 | strong | | TM |

**Table 4.1. Hybridoma cell lines selected for cloning of Fab fragments.**
Cell lines designated F1 were generated by immunization with the full length Rm1ClC, and F3 with the Rm1ClC transmembrane domain. The reaction of the antibodies in ELISA is qualitatively described. The antibodies recognizing sequence epitopes are labeled as Western positive (+). The binding to the full-length (FL), transmembrane (TM) and cytoplasmic domain (Dom) is indicated.

### 4.1.3 Material and methods

**Immunization of mice and fusion to myeloma cells.** The proteins were prepared and supplied to CEPower, where the immunization was carried out. Three mice were immunized with 300 µl of 300 µg/ml target protein. This procedure was repeated three times and 2 weeks after the last immunization the spleen cells were harvested. The spleen B-cell were fused to myeloma cells (ΔHGPRT) and cloned by limited dilution. The supernatants were tested for secreted antibodies by ELISA.

**ELISA.** 96 well MaxiSorp plates (Nunc) were coated with 100 µl of protein solution at the following concentration: full length Rm1ClC 30 µg/ml, transmembrane domain 20 µg/ml and cytoplasmic domain 7 µg/ml. The coating was carried out by vigorous shaking for 1 h at room temperature. Buffer B (10 mM Tris-HCl pH 7.4, 150 mM NaCl, 5 mM DM) was used in all steps. The plate was blocked for 1 h with 250 µl/well of buffer B + 5 % BSA (w/V). After blocking, 100 µl of hybridoma supernatants (1:1 mixed with 2 x buffer B containing 4 % BSA) were incubated for 1h followed by 3 washing steps for 5 minutes with buffer B (200 µl/well). For testing the polyclonal sera, 100 µl serum in serial dilutions of 1:100 to 1:800.000 in buffer B including 2 % BSA were used. The samples were subsequently incubated with 100 µl of goat anti-mouse IgG-HRP antibody (Roche) diluted to 1:2000 in buffer B including 2 % BSA for 1h and followed by 3 washing steps. The signal was developed with 100 µl of TMB substrate (Sigma) and the reaction was stopped by addition of 100 µl 1 M HCl. The TMB substrate consisted of 1 mM TMB, 3 mM $H_2O_2$, 25 mM citric acid, 50 mM $Na_2HPO_4$, pH 5. The signal was measured in the Tecan GENios Microplate Reader by recording absorbance at 450 nm.

**Western blot for antibody characterization.** A standard protocol for Western blots was used except for the following differences. The SDS-PA gels were prepared with only one small well for loading the marker and one big well spread over the entire gel for loading the protein sample. After electroblotting and blocking to prevent non-specific binding, the membrane was fixed in the Mini-PROTEAN II multiscreen apparatus (Bio-Rad), which allowed screening of 20 different antibodies per gel without having to cut the membrane into small strips. The goat anti-mouse IgG-HRP antibody (1:5000) was used as a secondary antibody according to the manufacturer's instructions (Roche).

## 4.2 Cloning and expression of different constructs of the ClC-1 cytoplasmic domain

ClC-1 is a chloride channel in mammalian skeletal muscles responsible for the maintainance of the resting membrane potential (Aromataris and Rychkov 2006). A reduction of the sarcolemal chloride conduction due to mutations in the ClC-1 gene results in myotonia, a disease that is accompanied by muscle hyperexcitability. Several studies in the last few years showed that ClC-1 can be inhibited by intracellular ATP, which acts on the common gate and shifts the voltage dependent activation to more positive potentials (Bennetts, Rychkov et al. 2005; Tseng, Bennetts et al. 2007). A more detailed analysis revealed a dependence of the ATP inhibition on the cytoplasmic pH and the redox state of the channel (Zhang, Tseng et al. 2008). The results indicate that the inhibition results from a direct interaction of ATP and the cytoplasmic domain of ClC-1, similar to the previously discovered nucleotide binding on the cytoplasmic domain of ClC-5 (Meyer, Savaresi et al. 2007). To reveal the structural basis for ATP regulation in ClC-1, I have investigated the expression of its cytoplasmic domain. In order to perform crystallization and co-crystallization trials with nucleotides, the cytoplasmic domain of ClC-1 needed to be produced in a soluble state in sufficient amounts. However, already this first step proved to be challenging, since the domain was expressed at the low levels mostly in inclusion bodies.

The cytoplasmic domain of ClC-1 channel is 398 residues long. Its unusually large size of is due to a long linker region connecting the two CBS subdomains (around 150 aa) and a long C-peptide (around 110 aa). The linker and the C-peptide regions are predicted to be disordered (Figure 4.2.1), which is not a suitable property when aiming for crystallization. It appears likely that unstructured regions are responsible for the aggregation of the protein upon overexpression.

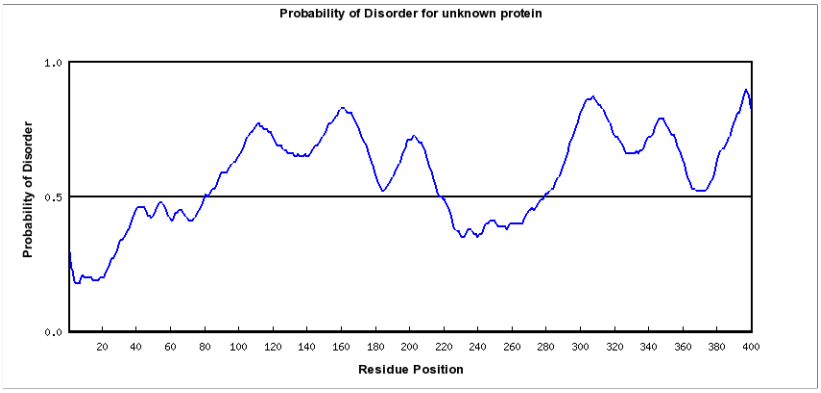



**Figure 4.2.1. Disorder prediction for the ClC-1 cytoplasmic domain.**
The residues having a high probability for being disordered belong to the linker region between the two CBS subdomains (81-215) and to the region of C-peptide (278-398) as predicted by the RONN server (Yang, Thomson et al. 2005).

In order to improve expression and to increase the chance of obtaining crystals of the ClC-1 cytoplasmic domain, expression constructs with reduced length of the flexible regions were designed. The first modification included a removal of the residues 675 – 797 in the linker region (Figure 4.2.2A). While the resulting construct named ClC-1-L1 domain showed a higher expression level a large fraction was insoluble similar as the full-length ClC-1 domain (Figure 4.2.2B). Starting from the construct with a considerably shortened liker region, two additional constructs with a shorter C-peptide were generated. In the construct ClC-1-L2 a stop codon was introduced after the residue 906 and in the construct ClC-1-L3 after the residue 879. In these constructs, the length of the C-peptide was similar as in ClC-0 and ClC-Ka, respectively. Both ClC-1-L2 and –L3 showed high expression level of soluble protein.

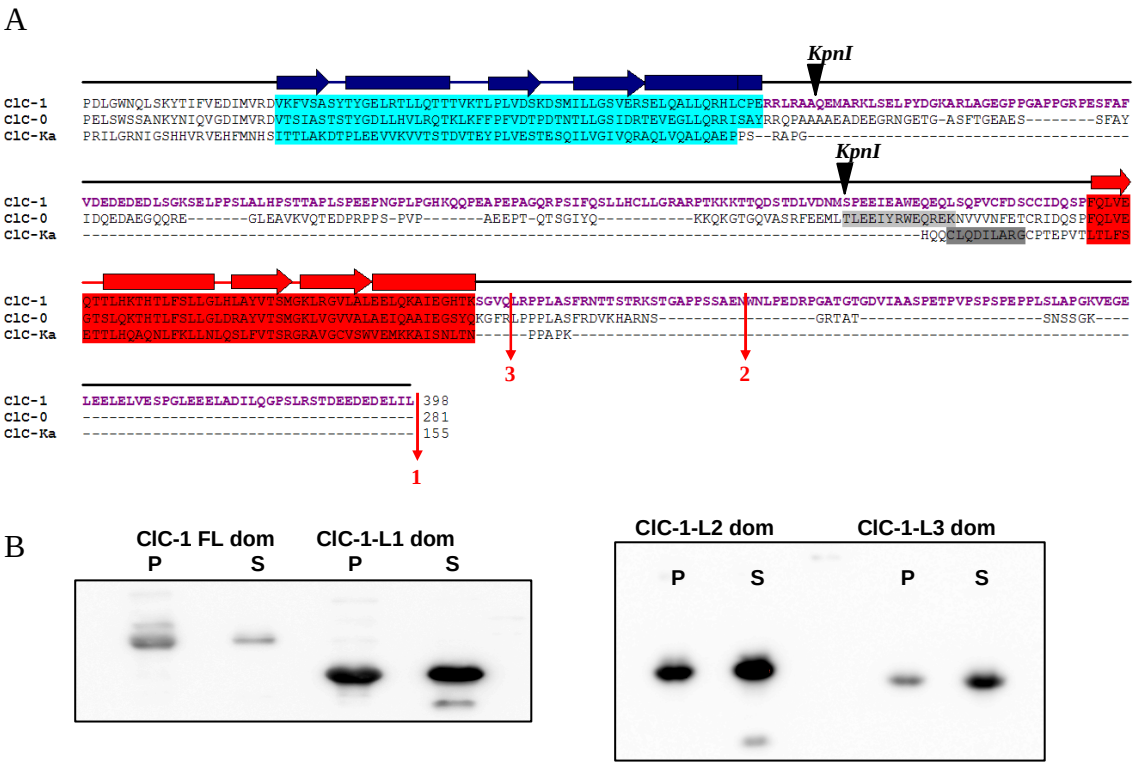



**Figure 4.2.2. Design of ClC-1 cytoplasmic domain constructs and the expression analysis.**
**(A)** Sequence alignment of the cytoplasmic domains of ClC-1, ClC-0 and ClC-Ka. The secondary structure elements are based on the structure of the ClC-Ka domain. CBS1 subdomain is colored in blue, CBS2 in red. Residues building the α-helix in the linker regions of ClC-0 and ClC-Ka are marked in grey. The construct ClC-1-L1 was generated by introducing two *KpnI* restriction sites at the presented positions, cutting out the gene fragment and subsequent religation. The constructs ClC-1-L2 and -L3 were generated by introducing stop codons at the positions labeled with red arrows in the construct ClC-1-L1.
**(B)** Expression/solubility test for different CLC-1 domain constructs. The insoluble fraction is labeled with P, soluble with S. All three shorter construct show better expression level and solubility compared to the full length ClC-1 cytoplasmic domain.

Purification of all three constructs yielded a high amount of monodisperse protein as illustrated in Figure 4.2.3. For crystallization trials only the construct ClC-1-L3 was

used, since it appeared to be the most promising one because of its compact shape. Extensive crystallization screening was performed at two different temperatures (4 and 20°C), different protein concentrations and in the presence of ATP, ADP and AMP.

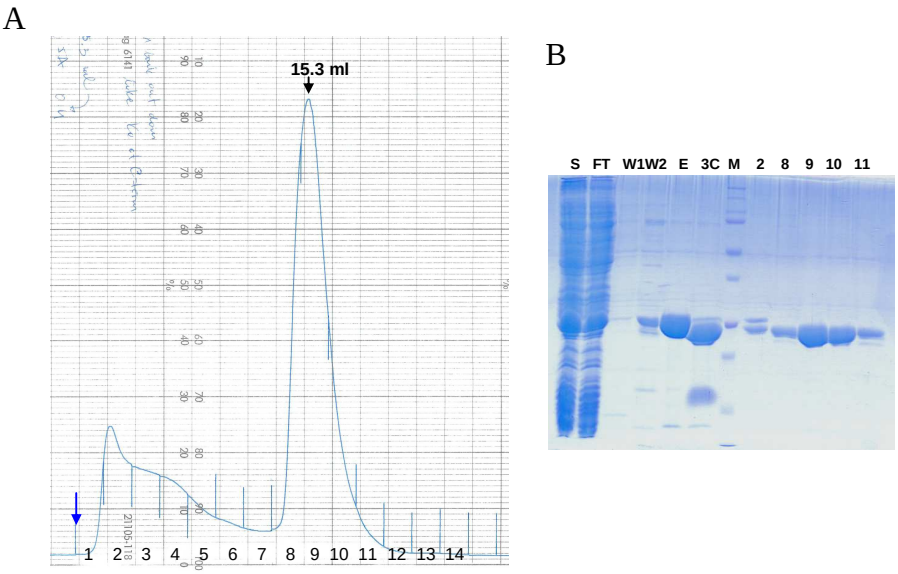



**Figure 4.2.3. Purification of the ClC-1-L3 domain construct.**
(A) Size exclusion chromatography on S200 column. The protein is monodisperse and elutes at 15.3 ml. The blue arrow designates the void volume. (B) SDS-PAGE analysis of the purification steps. S - cell lysate; FT – flow trough after metal affinity column; W1 and W2 – wash of metal affinity resin with 0 and 20 mM imidazole; E – fraction eluted from IMAC; 3C – protein cleaved with the 3C protease to remove the His-tag; 2, 8, 9, 10 and 11 – fractions from the SEC as labeled in A.

Unfortunately, none of the conditions contained crystals. It is difficult to assign the cause for this behavior but it could be that the designed constructs do not contain the optimal boundaries and that other constructs with variations in the linker and C-peptide region could still crystallize. Yet, even in other constructs the inherent conformational flexibility of the protein might still exist and prevent crystal formation.

Besides crystallization screening, the expressed ClC-1 domain constructs also allowed biochemical analyses. In order to test the oligomeric organization of the ClC-1 domain, AUC velocity experiments were performed. The sedimentation analysis at different protein concentration revealed a monomeric protein (data not shown), an unexpected feature when considering the assembly of the other eukaryotic ClC CBS domains. However, since it was not the native cytoplasmic domain of ClC-1, which has lead to this observation it will not be further discussed. In order to test whether ClC-1 cytoplasmic domain binds nucleotides, an isothermal titration calorimetry (ITC) experiment was performed. ITC measures the heat that is released upon ligand binding to a macromolecule and in that way allows to quantify the enthalpy and binding affinity. The titration of ATP to the ClC-1 domain is presented in Figure 4.2.4A. The obtained results suggest very weak binding of ATP to the ClC-1-L2 domain construct with the $K_d$ value of

about 400 µM. The affinity is, thus, around 15 times lower than the affinity of the ClC-5 domain for ATP determined by the same method (Figure 4.2.4B). This result also has to be taken with caution as it was not obtained on the native cytoplasmic domain of ClC-1.

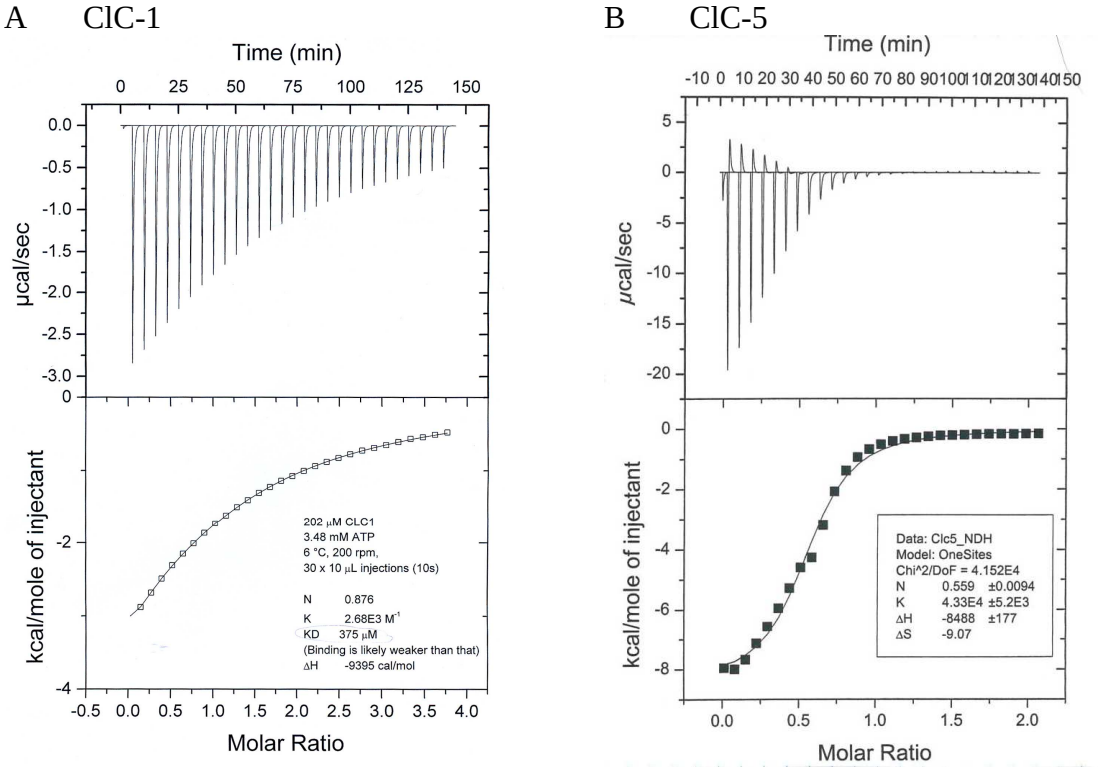



**Figure 4.2.4. Results of the ITC experiments for the ClC-1 (A) and ClC-5 (B) cytoplasmic domains.** Top: Raw data of the titration of ATP to the protein. Each peak shows the heat released by serial injections of ATP. Bottom: The total heat accumulated up to injection i is normalized to the total ATP concentration at step i and plotted against the ratio of total ATP concentration at step i to total protein concentration. From the sigmoid curve, ΔH per mol complex can be calculated. The solid line is the calculated curve using the best-fit parameters. Model: one binding site per protein molecule. The best values for stoichiometry (N), binding constant (K), enthalpy (H) and entropy (S) are also shown. The obtained $K_d$ values are: 375 µM for the ClC-1-L2 domain and 25 µM for the ClC-5 domain.

### 4.2.2 Patch clamp experiments on the ClC-1 channel expressed in *Xenopus* oocytes

In parallel with the biochemical experiments on the ClC-1 cytoplasmic domain, ClC-1 was expressed in *Xenopus leavis* oocytes and analyzed by electrophysiology. In order to study the functional effect of ATP binding to the cytoplasmic domain of ClC-1 excised inside-out patch clamp experiments were performed. The experiments were carried out similar as described by Tseng, Bennetts et al. 2007 and the obtained results are summarized in Figure 4.2.5. The current mediated by ClC-1 upon applying the voltage protocol A was recorded and analyzed. At pH 7.4 the channel is almost completely open at all positive voltages ($V_{1/2}$ = -39 mV). When lowering the pH from 7.4 to 6.2 the open probability shifts to

negative potentials. Upon addition of ATP to the intracellular solution at pH 6.2 in most of the cases no change in the open probability curve was observed. However, when excising the patch in solutions containing 1 mM ATP at pH 6.2 a pronounced effect of ATP was observed as shown in Figure 4.2.5. The inhibitory effect of ATP manifests itself in the shift of the open probability to more positive potentials. However, this effect was usually not reversible and vanished after some minutes. The explanation for this behavior was recently reported (Zhang, Tseng et al. 2008) and it is related with the oxidation of ClC-1 that seems to play an important role in the channel function. The inhibitory effect of ATP at low pH has not been observed under oxidizing conditions. The oxidation of the channel happens with time, since the reducing compounds from the cytoplasm are diffusing away.

In summary, although the calorimetric and electrophysiological analysis of ClC-1 revealed interesting features of this channel which hint at a possible interaction with ATP, definitive conclusions were not reached. Future investigations that take newly discovered features into account including the pH and oxidation dependence of ligand regulation might ultimately be more successful in reaching conclusive information on this important regulatory mechanism.

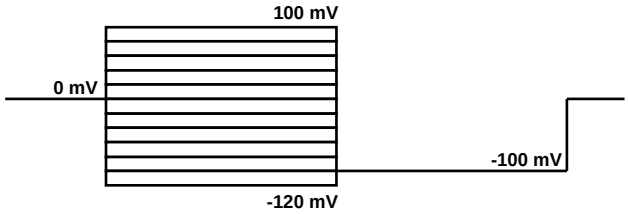
**Figure 4.2.5. Effect of pH and cytoplasmic ATP on ClC-1.**

**(A)** Voltage protocol A: The membrane potential was stepped from the 0 mV holding potential to various voltages ranging from +100 to -120 mV in (-20 mV steps) for 300 ms, followed by a step to -100 mV for 300 ms.

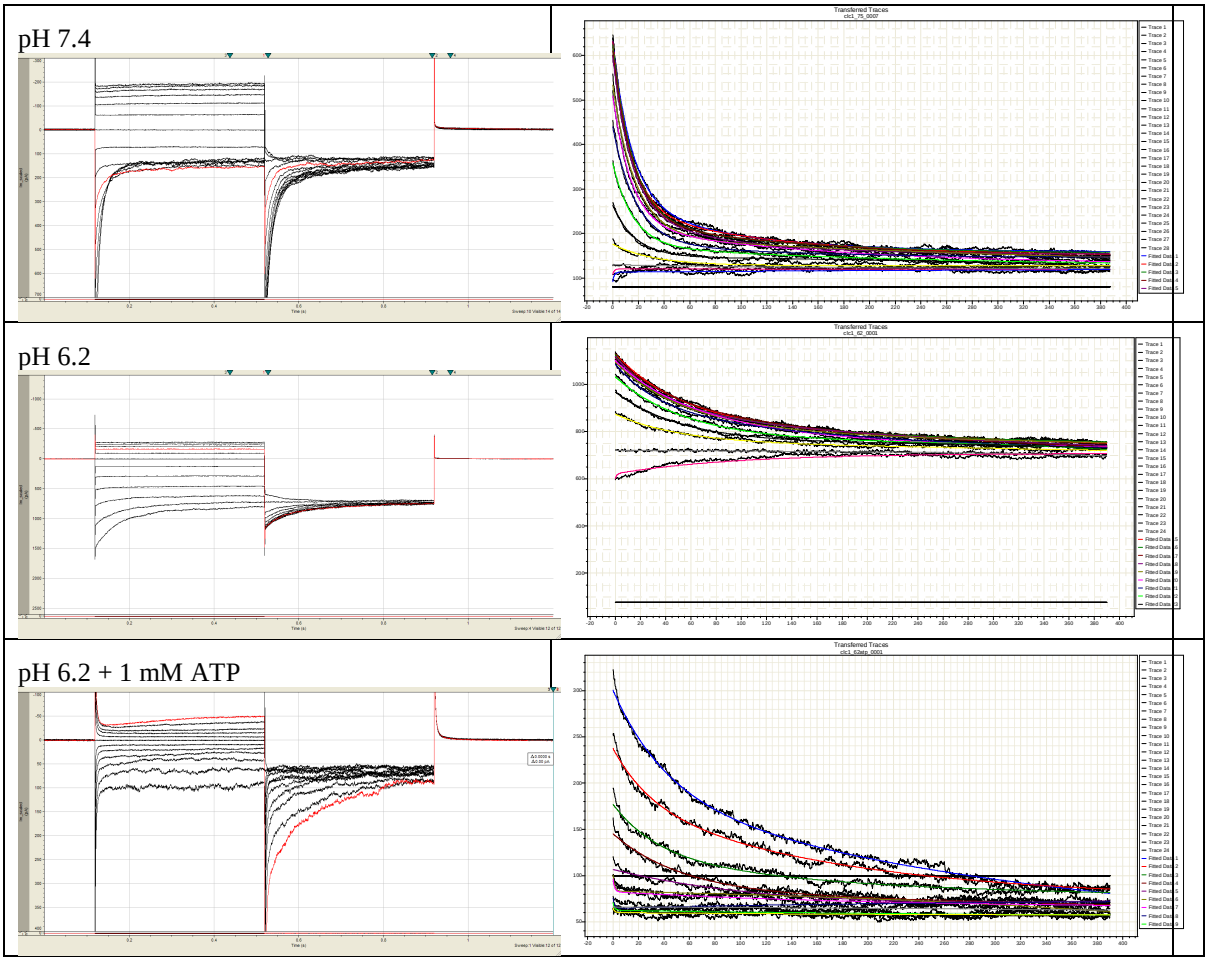
**(B)** Left: Current traces obtained by application of protocol A at the indicated conditions. Right: Fit of tail currents to a double exponential function with the program Clampfit allows to determine the initial tail current at each voltage.

**(C)** The normalized initial tail current is the product of the open probability of the fast and the common gate.

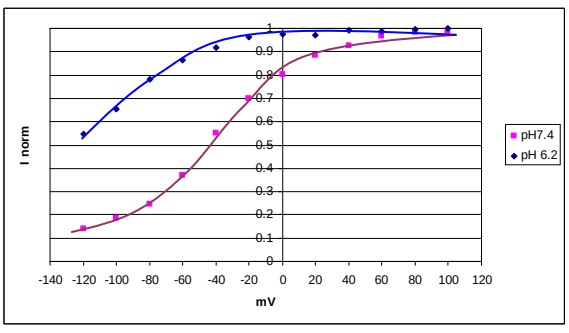
A

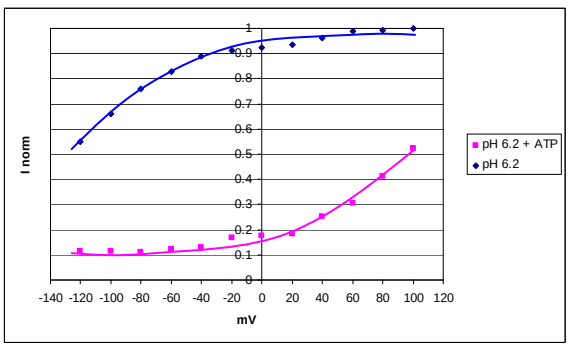

100 mV
0 mV
-100 mV
-120 mV

B


pH 7.4
pH 6.2
pH 6.2 + 1 mM ATP

C


I norm
mV
pH7.4
pH 6.2


I norm
mV
pH 6.2 + ATP
pH 6.2

### 4.2.3 Material and methods

**Cloning and expression of ClC-1 domain constructs.** Fragments of ClC-1 corresponding to desired constructs were amplified by PCR and cloned into the modified pET28b+ vector (Novagen) to generate fusions with the C-terminal 3C protease cutting site and a $His_6$-tag. The linker region was removed by introducing two *Kpn*I sites at specified positions with the QuikChange method (Stratagene), followed by digestion, elution and relegation of the ClC-1 vector. For expression the *E. coli* BL21 (DE3) strain was used. The cells were grown in LB medium containing 50 µg/ml kanamycin at 37°C to an $OD_{600}$ of 1.2 and induced by addition of 0.5 mM isopropyl-D-thiogalactopyranoside (IPTG) at 20°C over night. The protein was purified by standard procedures.

**ITC.** ITC experiments were performed at 6°C with the following concentrations: for ClC-1 200 µM protein and 3.5 mM ATP and for ClC-5 500 µM protein and 5 mM ATP. Before the measurement both the protein and ATP were dialyzed against 10 mM Tris-HCl, 150 mM NaCl to ensure the same buffer composition for both components.

**Patch clamp recordings.** The human ClC-1 channel in the pTLN vector was used for mRNA synthesis using SP6 mMessage mMachine kit (Ambion). For expression, 20 ng of ClC-1 mRNA was injected into defolliculated *Xenopus leavis* oocytes. From 3-5 days after injection, excised inside-out patch recordings were performed, using the Axopatch 200B amplifier and a Digidata 1440A board controlled by pClamp8 software (Axon Instrumets, Inc./Molecular Devices). The glass electrodes had a tip size of 1-2 µm and a resistance of 4-5 MΩ and were filled with a pipette (extracellular) solution containing: 120 mM NMG-Cl, 1 $MgCl_2$, 10 mM HEPES, 1 mM EGTA, pH 7.4. The bath (intracellular) solution had the same ionic components with the pH adjusted to 7.4 or 6.2. Working solutions of ATP were freshly made by mixing the bath solution and 100 mM ATP stock solution and subsequent adjustment of the pH.

Macroscopic chloride currents were recorded upon applying the voltage protocol A presented in Figure 4.2.5A. The initial value of the tail current was determined by fitting a double exponential function to the tail current using the program Clampfit. The initial current of each trace was normalized to the maximal value of the initial tail current obtained following the most positive test voltage in the absence of ATP. The normalized initial tail current obtained with this protocol represents the product of the open probability of the fast gate and of the common gate at the preceding test voltage.

## 4.3 Expression of eukaryotic ClC homologs in *Pichia pastoris*

For a comprehensive understanding of the mechanisms of ion conduction and gating in eukaryotic ClC proteins, detailed structural and functional investigations have to be carried out on full length eukaryotic family members. The current bottleneck for the investigation of those integral membrane proteins is the difficulty to express them in sufficient quantities. Therefore, one of the main challenges during my PhD thesis was the establishment of an expression system for eukaryotic ClCs. Since different mammalian membrane proteins, including G-protein coupled receptors and ion channels, have been successfully expressed in the methylotrophic yeast *Pichia pastoris* (Macauley-Patrick, Fazenda et al. 2005), this system was chosen for expression of eukaryotic ClC channels and transporters.

Several yeast ClC homologs (summarized in Table 4.2), the channel ClC-0 from *Torpedo marmorata* and the human proteins ClC-1, ClC-2, ClC-Ka and ClC-5 were selected for cloning and expression in *P. pastoris*. The applied cloning and expression protocols were mostly based on the Invitrogen manuals. All genes were cloned in the vector pPICZB resulting in C-terminal fusion with *c-myc* and $His_6$-tags. After preparation of recombinant plasmids in *E. coli*, the transformation into *P. pastoris* and expression tests were performed. Unfortunately, none of the selected ClC homologs showed expression, even though the integration into the *Pichia* genome was confirmed.

| **Strain with the strain code** | **Gene** | **Identity to ClC-5 [%]** | **Length [aa]** | **Mw [kD]** |
|---|---|---|---|---|
| *Pichia pastoris* GS115 | PAS_chr2-1_0471 *Part of this gene* | 33 | 650 | 71.9 |
| *Pichia pastoris* GS115 | - | 30 | 936 | 104.5 |
| *Saccharomyces cerevisiae* | C1Q_03652 | 30 | 779 | 87.6 |
| *Schizosaccharomyces pombe* 972h- | SPBC887.02 | 27 | 667 | 73.3 |
| *Schizosaccharomyces pombe* 972h- | SPBC19C7.11 | 29 | 812 | 90.5 |
| *Yarrowia lypolitica* CLIB122 | YALI0E09015g | 38 | 772 | 85.3 |

**Table 4.2. Yeast ClC Homologs.**
List of 6 ClC homologs from yeasts selected for cloning. Identity to ClC-5 was determined by pairwise sequence alignments implemented in ClustalW.

Following unsuccessful expression of different ClC homologs, I pursued with another approach. ClC-0 was selected as a target gene and cloned with three different N-

terminal tags: an $His_8$-tag (MAHHHHHHHH), a Flag-tag (MADYKDDDDK) and a StrepII-tag (MAWSHPQFEK). In all three constructs the tag was followed by a cleavage site for the 3C protease. The DNA sequence corresponding to the N-terminal tag and protease recognition site was optimized for *Pichia* codon usage. The N-terminally modified ClC-0 constructs were cloned into pPICZB vector and standard expression procedures were applied. Expression was detected only in the construct containing a N-terminal $His_8$-tag (N-$His_8$-ClC-0).

The expression conditions were optimized by varying media, temperature and time of induction. Short expression for 6-8 h after induction at 30°C proved to give the highest yields. After that time, the protein was not detectable by Western blot, probably due to degradation.

The extraction of N-$His_8$-ClC-0 from the membrane was tested on a small scale with a set of several detergents including dodecyl-β-D-maltoside (DDM), tetradecyl-β-D-maltoside (TDM) and Fos-choline-12 (Fos12). The protein was almost completely extractable with Fos12 and partially with DDM and TDM (Figure 4.3.1).

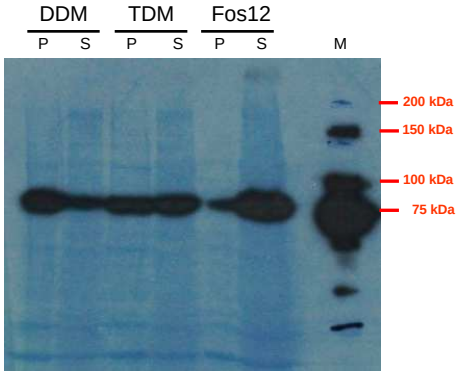



**Figure 4.3.1. Detergent extraction test for N-$His_8$-ClC-0.**
The membranes were solubilized with 2 % DDM, 2 % TDM and 1 % Fos12 respectively. The samples were centrifuged at high speed (100,000xg) to remove insoluble parts. The solubilized (labeled S) and not solubilized fractions (labeled P) were analyzed by Western blot to determine the amount of extracted protein. ClC-0 is extractable with Fos12, and partially with DDM and TDM.

In order to assess the protein stability in different detergents, the size distribution was analyzed by size-exclusion chromatography. In a simple assay, the crude membrane extract from a small scale solubilization after centrifugation was applied on a Superdex200 column. The retention volume of the protein was determined by Western blot analysis of the elution fractions starting from the void volume of 7 ml to 13 ml (Figure 4.3.2). The results indicated promising behavior of N-$His_8$-ClC-0, since the protein was eluting both in DDM and Fos-12 at 10 - 10.5 ml, corresponding to an appropriate molecular weight of the dimeric ClC-0. Since DDM is a milder detergent than Fos-12, it was selected for extraction and purification in the subsequent experiments.

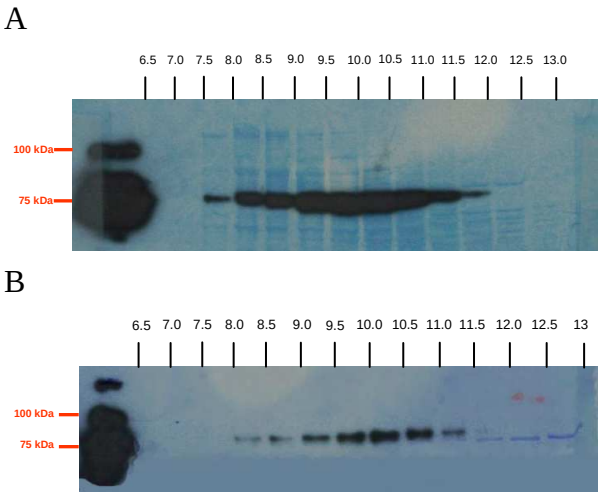



**Figure 4.3.2. Size exclusion chromatography of N-His$_8$-ClC-0.** The membranes of *P. pastoris* cells expressing ClC-0 were solubilized by Fos12 **(A)** and DDM **(B)**. The solubilized fraction was injected on Superdex200 column. The eluted fractions were analyzed by Western blot with an antibody recognizing a His tag. The void volume of the column is 7 ml. In both Fos12 and DDM, the elution volume of ClC-0 is around 10 ml, which would correspond to the approximate molecular weight of a dimer surrounded by a detergent micelle. The elution volumes matching the analyzed elution fractions are indicated.

The large-scale purification of N-His$_8$-ClC-0 was performed applying usual protocols for membrane protein purification. The protein was extracted from the membrane with DDM. Additional purification steps included metal affinity and size exclusion chromatography. The N-His$_8$-ClC-0 construct showed binding to the metal affinity resin and was eluted with increasing imidazole concentration. However, the eluted fraction was not of satisfying purity and the protein was detected only by Western blot (Figure 4.3.3). Eluted fractions containing ClC-0 were concentrated and injected on a Superdex200 column. The Western blot analysis of the size exclusion fractions revealed that most of the protein was eluting at the void volume due to the formation of high-molecular aggregates during the purification.

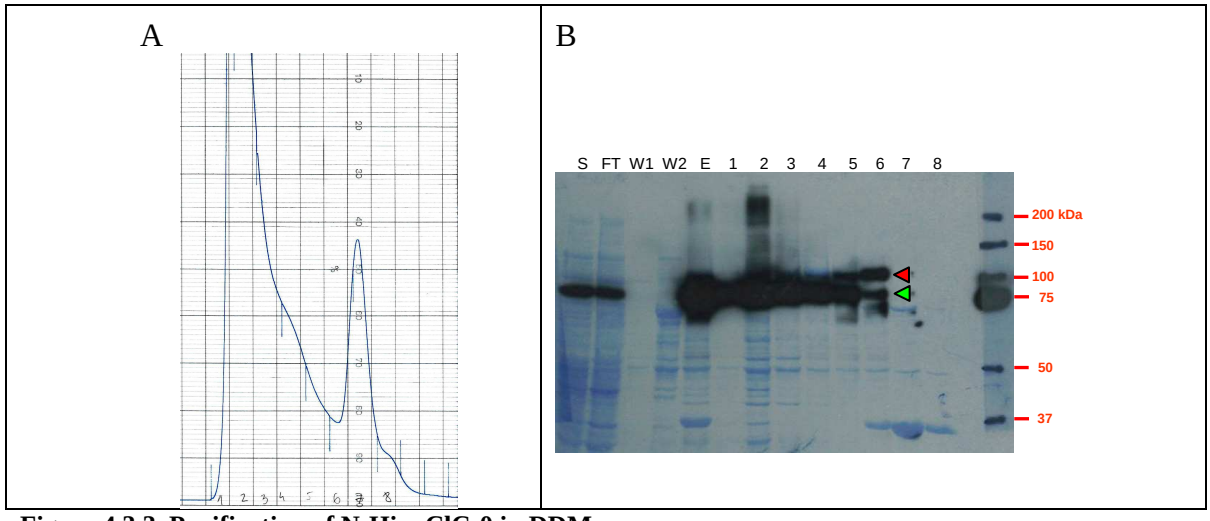



**Figure 4.3.3. Purification of N-His$_8$-ClC-0 in DDM.**
**(A)** Size exclusion chromatography on S200 column. The amount of aggregated proteins is high. **(B)** Western blot analysis of the purification steps: S - cell lysate; FT – flow trough after metal affinity column; W1 and W2 – wash of metal affinity resin with 0 and 20 mM imidazole; E – fraction eluted from IMAC; 1,2, 3, 4, 5, 6, 7 and 8 – fractions from the SEC as labeled in A. The reason for the different migration of the two bands in fraction 6 is unknown, but they might present glycosylated (red triangle) and not-glycosylated ClC-0 (green triangle).

In summary, expression screening of several eukaryotic ClC proteins have led to the identification of a potential candidate for continuing investigations. The identified ClC-0 construct showed promising extraction and purification behavior when tested on small scale, but large scale expression and purification revealed the tendency for aggregation making it unsuitable for crystallographic studies in its current state.

### 4.3.1 Material and methods

**Bacterial and yeast strains and media.** For cloning and production of recombinant plasmids the *Escherichia coli* strain XL1blue was used. *E. coli* transformants were selected on low salt LB plates with 25 µg/ml zeocin. The protease deficient *P. pastoris* strain SMD1163 (*his4, pep4, prb*) (Invitrogen) was used for expression. This strain was generally grown on YPD medium. For expression studies a combination of YPD and YPM medium or a combination of the buffered glycerol-complex medium (BMGY) and buffered methanol-complex medium (BMMY) was used. *P. pastoris* transformants were selected on YPDS plates with 100 µg/ml zeocin. The composition of all these media was as indicated in the Invitrogen instruction manuals (*Pichia* Expression Kit and pPICZ A, B and C).

**Construction of the expression plasmids and transformation.** The target genes were amplified from genomic or plasmid DNA and cloned into the plasmid pPICZB. After transformation of ligation mixtures into *E. coli*, plasmids were isolated and the inserted genes sequenced. Approximately 10 µg of recombinant plasmids were linearized with *Pme*I and used to transform competent cells of *P. pastoris* by electroporation. The transformation was done according to the protocol in Invitrogen manual *Pichia* Expression Kit (page 59). Transformants were selected on YPDS plates with 100 µg/ml zeocin after 2 days incubation at 30°C.

**Expression tests**. 10 randomly picked transformants were grown in 4 ml of YPD or BMGY medium supplemented with 25 µg/ml zeocin overnight at 30°C under vigorous shaking. The cultures were pelleted and washed in 4 ml of 0.1 M K-phosphate buffer pH 6. The cell pellets were resuspended in 4 ml of induction medium containing 0.5 % methanol (YPM or BMMY). The expression was carried out either at 30°C or 23°C. For the time course analysis of expression induction time was varied from 4 to 24 h. After expression, the amount of cells that correspond to $OD_{600}$=4 was pelleted and resuspended in 400 µl ice-cold breaking buffer (50 mM Na-phosphate pH 7.4, 1 mM PMSF, 5 % glycerol and 1 mM EDTA). The cells were broken by addition of glass beads (300 mg, 0.6 mm diameter) in a Fast prep device (Bio101) for two times 30 s at force 6.5 with 5 min cooling on ice

between the runs. The lysates were separated by SDS-PAGE and analyzed for expression by Western blot.

**Large scale expression of ClC-0.** For large scale expression of ClC-0 4 l culture was used. The cells were grown in YPD medium over night at 30°C, washed in 0.1 M K-phosphate buffer pH 6 to remove glucose and resuspended in BMMY medium to final $OD_{600}$=1. The expression was carried out at 30°C for 6-8 h. Following expression, the cells were harvested by centrifugation and resuspended in 60 ml of breaking buffer. The suspension was distributed into 4x50 ml Falcon tubes and a total amount of 60 ml of glass beads was added. The cells were broken in a Fast prep device for two times 30 s at force 6.5. The lysate was cleared by low-spin centrifugation (30 min, 10,000 x g, +4ºC). The membranes were pelleted from the supernatant by high-spin centrifugation (1.5 h, 100,000 x g, +4ºC). The protein was extracted from the membrane by 2 % DDM and purification performed applying standard procedures.

# 5 Bibliography

Accardi, A. and C. Miller (2004). "Secondary active transport mediated by a prokaryotic homologue of ClC Cl- channels." Nature **427**(6977): 803-807.

Accardi, A. and M. Pusch (2000). "Fast and slow gating relaxations in the muscle chloride channel CLC-1." J Gen Physiol **116**(3): 433-444.

Accardi, A., M. Walden, et al. (2005). "Separate ion pathways in a Cl-/H+ exchanger." J Gen Physiol **126**(6): 563-570.

Adams, P. D., R. W. Grosse-Kunstleve, et al. (2002). "PHENIX: building new software for automated crystallographic structure determination." Acta Crystallogr D Biol Crystallogr **58**(Pt 11): 1948-1954.

Alioth, S., S. Meyer, et al. (2007). "The cytoplasmic domain of the chloride channel ClC-0: structural and dynamic characterization of flexible regions." J Mol Biol **369**(5): 1163-1169.

Aromataris, E. C. and G. Y. Rychkov (2006). "ClC-1 chloride channel: Matching its properties to a role in skeletal muscle." Clin Exp Pharmacol Physiol **33**(11): 1118-1123.

Bailey, S. (1994). "The Ccp4 suite - programs for protein crystallography." Acta Crystallographica Section D-Biological Crystallography **50**(760-763).

Barth, H. G., B. E. Boyes, et al. (1994). "Size exclusion chromatography." Anal Chem **66**(12): 595R-620R.

Bateman, A. (1997). "The structure of a domain common to archaebacteria and the homocystinuria disease protein." Trends Biochem Sci **22**(1): 12-13.

Bennetts, B., M. W. Parker, et al. (2007). "Inhibition of skeletal muscle ClC-1 chloride channels by low intracellular pH and ATP." J Biol Chem **282**(45): 32780-32791.

Bennetts, B., M. L. Roberts, et al. (2001). "Temperature dependence of human muscle ClC-1 chloride channel." J Physiol **535**(Pt 1): 83-93.

Bennetts, B., G. Y. Rychkov, et al. (2005). "Cytoplasmic ATP-sensing domains regulate gating of skeletal muscle ClC-1 chloride channels." J Biol Chem **280**(37): 32452-32458.

Bernard, P., P. Gabant, et al. (1994). "Positive-selection vectors using the F plasmid ccdB killer gene." Gene **148**(1): 71-74.

Biemans-Oldehinkel, E., N. A. Mahmood, et al. (2006). "A sensor for intracellular ionic strength." Proc Natl Acad Sci U S A **103**(28): 10624-10629.

Binz, H. K., P. Amstutz, et al. (2004). "High-affinity binders selected from designed ankyrin repeat protein libraries." Nat Biotechnol **22**(5): 575-582.

Bosl, M. R., V. Stein, et al. (2001). "Male germ cells and photoreceptors, both dependent on close cell-cell interactions, degenerate upon ClC-2 Cl(-) channel disruption." EMBO J **20**(6): 1289-1299.

Brandt, S. and T. J. Jentsch (1995). "ClC-6 and ClC-7 are two novel broadly expressed members of the CLC chloride channel family." FEBS Lett **377**(1): 15-20.

Bykova, E. A., X. D. Zhang, et al. (2006). "Large movement in the C terminus of CLC-0 chloride channel during slow gating." Nat Struct Mol Biol **13**(12): 1115-1119.

Doyle, D. A., J. Morais Cabral, et al. (1998). "The structure of the potassium channel: molecular basis of K+ conduction and selectivity." Science **280**(5360): 69-77.

Drew, D., M. Lerch, et al. (2006). "Optimization of membrane protein overexpression and purification using GFP fusions." Nat Methods **3**(4): 303-313.

Drew, D. E., G. von Heijne, et al. (2001). "Green fluorescent protein as an indicator to monitor membrane protein overexpression in Escherichia coli." FEBS Lett **507**(2): 220-224.

Duffield, M., G. Rychkov, et al. (2003). "Involvement of helices at the dimer interface in ClC-1 common gating." J Gen Physiol **121**(2): 149-161.

Dutzler, R., E. B. Campbell, et al. (2002). "X-ray structure of a ClC chloride channel at 3.0 A reveals the molecular basis of anion selectivity." Nature **415**(6869): 287-294.

Dutzler, R., E. B. Campbell, et al. (2003). "Gating the selectivity filter in ClC chloride channels." Science **300**(5616): 108-112.

Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallogr D Biol Crystallogr **60**(Pt 12 Pt 1): 2126-2132.

Estevez, R., T. Boettger, et al. (2001). "Barttin is a Cl- channel beta-subunit crucial for renal Cl- reabsorption and inner ear K+ secretion." Nature **414**(6863): 558-561.

Estevez, R., M. Pusch, et al. (2004). "Functional and structural conservation of CBS domains from CLC chloride channels." J Physiol **557**(Pt 2): 363-378.

Fong, P., A. Rehfeldt, et al. (1998). "Determinants of slow gating in ClC-0, the voltage-gated chloride channel of Torpedo marmorata." Am J Physiol **274**(4 Pt 1): C966-973.

Gadsby, D. C. (2009). "Ion channels versus ion pumps: the principal difference, in principle." Nat Rev Mol Cell Biol **10**(5): 344-352.

Geertsma, E. R., M. Groeneveld, et al. (2008). "Quality control of overexpressed membrane proteins." Proc Natl Acad Sci U S A **105**(15): 5722-5727.

Gunther, W., A. Luchow, et al. (1998). "ClC-5, the chloride channel mutated in Dent's disease, colocalizes with the proton pump in endocytotically active kidney cells." Proc Natl Acad Sci U S A **95**(14): 8075-8080.

Hammon, J., D. V. Palanivelu, et al. (2009). "A green fluorescent protein screen for identification of well-expressed membrane proteins from a cohort of extremophilic organisms." Protein Sci **18**(1): 121-133.

Hattori, M., N. Iwase, et al. (2009). "Mg(2+)-dependent gating of bacterial MgtE channel underlies Mg(2+) homeostasis." EMBO J **28**(22): 3602-3612.

Hebert, S. C. (2003). "Bartter syndrome." Curr Opin Nephrol Hypertens **12**(5): 527-532.

Hilf, R. J. and R. Dutzler (2008). "X-ray structure of a prokaryotic pentameric ligand-gated ion channel." Nature **452**(7185): 375-379.

Hogbom, M., M. Eklund, et al. (2003). "Structural basis for recognition by an in vitro evolved affibody." Proc Natl Acad Sci U S A **100**(6): 3191-3196.

Ignoul, S. and J. Eggermont (2005). "CBS domains: structure, function, and pathology in human proteins." Am J Physiol Cell Physiol **289**(6): C1369-1378.

Jayaram, H., A. Accardi, et al. (2008). "Ion permeation through a $Cl^-$-selective channel designed from a CLC Cl-/H+ exchanger." Proc Natl Acad Sci U S A **105**(32): 11194-11199.

Jentsch, T. J. (2008). "CLC chloride channels and transporters: from genes to protein structure, pathology and physiology." Crit Rev Biochem Mol Biol **43**(1): 3-36.

Jentsch, T. J., V. Stein, et al. (2002). "Molecular structure and physiological function of chloride channels." Physiol Rev **82**(2): 503-568.

Jentsch, T. J., K. Steinmeyer, et al. (1990). "Primary structure of Torpedo marmorata chloride channel isolated by expression cloning in Xenopus oocytes." Nature **348**(6301): 510-514.

Kabsch, W. (1993). "Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants." Journal of Applied Crystallography **26**: 795-800.

Kawate, T. and E. Gouaux (2006). "Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins." Structure **14**(4): 673-681.

Kieferle, S., P. Fong, et al. (1994). "Two highly homologous members of the ClC chloride channel family in both rat and human kidney." Proc Natl Acad Sci U S A **91**(15): 6943-6947.

Koch, M. C., K. Steinmeyer, et al. (1992). "The skeletal muscle chloride channel in dominant and recessive human myotonia." Science **257**(5071): 797-800.

Koide, S. (2009). "Engineering of recombinant crystallization chaperones." Curr Opin Struct Biol **19**(4): 449-457.

Kornak, U., D. Kasper, et al. (2001). "Loss of the ClC-7 chloride channel leads to osteopetrosis in mice and man." Cell **104**(2): 205-215.

Krissinel, E. and K. Henrick (2004). "Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions." Acta Crystallogr D Biol Crystallogr **60**(Pt 12 Pt 1): 2256-2268.

Krissinel, E. and K. Henrick (2007). "Inference of macromolecular assemblies from crystalline state." J Mol Biol **372**(3): 774-797.

Krissinel, E. a. H., K. (2004). "Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions." Acta Crystallogr. **D60**: 2256-2268.

Lange, P. F., L. Wartosch, et al. (2006). "ClC-7 requires Ostm1 as a beta-subunit to support bone resorption and lysosomal function." Nature **440**(7081): 220-223.

Langer, G., S. X. Cohen, et al. (2008). "Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7." Nat Protoc **3**(7): 1171-1179.

Lisal, J. and M. Maduke (2008). "The ClC-0 chloride channel is a 'broken' Cl-/H+ antiporter." Nat Struct Mol Biol **15**(8): 805-810.

Lobet, S. and R. Dutzler (2006). "Ion-binding properties of the ClC chloride selectivity filter." EMBO J **25**(1): 24-33.

Locher, K. P., A. T. Lee, et al. (2002). "The E. coli BtuCD structure: a framework for ABC transporter architecture and mechanism." Science **296**(5570): 1091-1098.

Long, S. B., E. B. Campbell, et al. (2005). "Voltage sensor of Kv1.2: structural basis of electromechanical coupling." Science **309**(5736): 903-908.

Looser, T. (2008). "Structural characterization of the C-terminal domain of the prokaryotic ClC homologue from *Picrophilus torridus* " Master thesis. University of Zurich.

Lucas, M., J. A. Encinar, et al. (2009). "Binding of S-Methyl-5'-Thioadenosine and S-Adenosyl-l-Methionine to Protein MJ0100 Triggers an Open-to-Closed Conformational Change in Its CBS Motif Pair." J Mol Biol.

Luecke, H., B. Schobert, et al. (2001). "Crystal structure of sensory rhodopsin II at 2.4 angstroms: insights into color tuning and transducer interaction." Science **293**(5534): 1499-1503.

Macauley-Patrick, S., M. L. Fazenda, et al. (2005). "Heterologous protein production using the Pichia pastoris expression system." Yeast **22**(4): 249-270.

Maduke, M., D. J. Pheasant, et al. (1999). "High-level expression, functional reconstitution, and quaternary structure of a prokaryotic ClC-type chloride channel." J Gen Physiol **114**(5): 713-722.

Maduke, M., C. Williams, et al. (1998). "Formation of CLC-0 chloride channels from separated transmembrane and cytoplasmic domains." Biochemistry **37**(5): 1315-1321.

Markovic, S. and R. Dutzler (2007). "The structure of the cytoplasmic domain of the chloride channel ClC-Ka reveals a conserved interaction interface." Structure **15**(6): 715-725.

Matsumura, Y., S. Uchida, et al. (1999). "Overt nephrogenic diabetes insipidus in mice lacking the CLC-K1 chloride channel." Nat Genet **21**(1): 95-98.

Matthews, B. W. (1968). "Solvent content of protein crystals." J Mol Biol **33**(2): 491-497.

Matulef, K. and M. Maduke (2007). "The CLC 'chloride channel' family: revelations from prokaryotes." Mol Membr Biol **24**(5-6): 342-350.

McCoy, A. J., R. W. Grosse-Kunstleve, et al. (2007). "Phaser crystallographic software." J Appl Crystallogr **40**(Pt 4): 658-674.

Meyer, S. and R. Dutzler (2006). "Crystal structure of the cytoplasmic domain of the chloride channel ClC-0." Structure **14**(2): 299-307.

Meyer, S., S. Savaresi, et al. (2007). "Nucleotide recognition by the cytoplasmic domain of the human chloride transporter ClC-5." Nat Struct Mol Biol **14**(1): 60-67.

Middleton, R. E., D. J. Pheasant, et al. (1994). "Purification, reconstitution, and subunit composition of a voltage-gated chloride channel from Torpedo electroplax." Biochemistry **33**(45): 13189-13198.

Miller, C. (1982). "Open-state substructure of single chloride channels from Torpedo electroplax." Philos Trans R Soc Lond B Biol Sci **299**(1097): 401-411.

Miller, C. (2006). "ClC chloride channels viewed through a transporter lens." Nature **440**(7083): 484-489.

Miller, C. and M. M. White (1984). "Dimeric structure of single chloride channels from Torpedo electroplax." Proc Natl Acad Sci U S A **81**(9): 2772-2775.

Mohammad-Panah, R., R. Harrison, et al. (2003). "The chloride channel ClC-4 contributes to endosomal acidification and trafficking." J Biol Chem **278**(31): 29267-29277.

Nguitragool, W. and C. Miller (2007). "Inaugural Article: CLC Cl /H+ transporters constrained by covalent cross-linking." Proc Natl Acad Sci U S A **104**(52): 20659-20665.

P.D. Adams, R. W. G.-K., L.-W. Hung, T.R. Ioerger, A.J. McCoy, N.W. Moriarty, R.J. Read, J.C. Sacchettini, N.K. Sauter and T.C. Terwilliger (2002). "PHENIX: building new software for automated crystallographic structure determination." Acta Cryst. **D58**: 1948-1954.

Pedersen, T. H., F. de Paoli, et al. (2005). "Increased excitability of acidified skeletal muscle: role of chloride conductance." J Gen Physiol **125**(2): 237-246.

Picollo, A., M. Malvezzi, et al. (2009). "Basis of substrate binding and conservation of selectivity in the CLC family of channels and transporters." Nat Struct Mol Biol **16**(12): 1294-1301.

Picollo, A. and M. Pusch (2005). "Chloride/proton antiporter activity of mammalian CLC proteins ClC-4 and ClC-5." Nature **436**(7049): 420-423.

Poet, M., U. Kornak, et al. (2006). "Lysosomal storage disease upon disruption of the neuronal chloride transport protein ClC-6." Proc Natl Acad Sci U S A **103**(37): 13854-13859.

Prive, G. G. (2007). "Detergents for the stabilization and crystallization of membrane proteins." Methods **41**(4): 388-397.

Pusch, M., U. Ludewig, et al. (1997). "Temperature dependence of fast and slow gating relaxations of ClC-0 chloride channels." J Gen Physiol **109**(1): 105-116.

Pusch, M., U. Ludewig, et al. (1995). "Gating of the voltage-dependent chloride channel CIC-0 by the permeant anion." Nature **373**(6514): 527-531.

Pusch, M., K. Steinmeyer, et al. (1994). "Low single channel conductance of the major skeletal muscle chloride channel, ClC-1." Biophys J **66**(1): 149-152.

Roosild, T. P., K. T. Le, et al. (2004). "Cytoplasmic gatekeepers of K+-channel flux: a structural perspective." Trends Biochem Sci **29**(1): 39-45.

Rudolph, M. J., G. A. Amodeo, et al. (2007). "Structure of the Bateman2 domain of yeast Snf4: dimeric association and relevance for AMP binding." Structure **15**(1): 65-74.

Rychkov, G. Y., M. Pusch, et al. (1998). "Permeation and block of the skeletal muscle chloride channel, ClC-1, by foreign anions." J Gen Physiol **111**(5): 653-665.

Scheel, O., A. A. Zdebik, et al. (2005). "Voltage-dependent electrogenic chloride/proton exchange by endosomal CLC proteins." Nature **436**(7049): 424-427.

Schleif, R. (2003). "AraC protein: a love-hate relationship." Bioessays **25**(3): 274-282.

Schmidt-Rose, T. and T. J. Jentsch (1997). "Reconstitution of functional voltage-gated chloride channels from complementary fragments of CLC-1." J Biol Chem **272**(33): 20515-20521.

Schneider, T. R. and G. M. Sheldrick (2002). "Substructure solution with SHELXD." Acta Crystallogr D Biol Crystallogr **58**(Pt 10 Pt 2): 1772-1779.

Scholl, U., S. Hebeisen, et al. (2006). "Barttin modulates trafficking and function of ClC-K channels." Proc Natl Acad Sci U S A **103**(30): 11411-11416.

Schriever, A. M., T. Friedrich, et al. (1999). "CLC chloride channels in Caenorhabditis elegans." J Biol Chem **274**(48): 34238-34244.

Schuck, P. (2000). "Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling." Biophys J **78**(3): 1606-1619.

Schuck, P., M. A. Perugini, et al. (2002). "Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems." Biophys J **82**(2): 1096-1111.

Schwake, M., T. Friedrich, et al. (2001). "An internalization signal in ClC-5, an endosomal Cl-channel mutated in dent's disease." J Biol Chem **276**(15): 12049-12054.

Schwappach, B., S. Stobrawa, et al. (1998). "Golgi localization and functionally important domains in the NH2 and COOH terminus of the yeast CLC putative chloride channel Gef1p." J Biol Chem **273**(24): 15110-15118.

Scott, J. W., S. A. Hawley, et al. (2004). "CBS domains form energy-sensing modules whose binding of adenosine ligands is disrupted by disease mutations." J Clin Invest **113**(2): 274-284.

Sennhauser, G., P. Amstutz, et al. (2007). "Drug export pathway of multidrug exporter AcrB revealed by DARPin inhibitors." PLoS Biol **5**(1): e7.

Sheldrick , G. M. (2002). "Macromolecular phasing with SHELXE." Z. Kristallogr. **217**: 644-650.

Steinmeyer, K., C. Ortland, et al. (1991). "Primary structure and functional expression of a developmentally regulated skeletal muscle chloride channel." Nature **354**(6351): 301-304.

Steinmeyer, K., B. Schwappach, et al. (1995). "Cloning and functional expression of rat CLC-5, a chloride channel related to kidney disease." J Biol Chem **270**(52): 31172-31177.

Stobrawa, S. M., T. Breiderhoff, et al. (2001). "Disruption of ClC-3, a chloride channel expressed on synaptic vesicles, leads to a loss of the hippocampus." Neuron **29**(1): 185-196.

Strange, K. (2003). "From genes to integrative physiology: ion channel and transporter biology in Caenorhabditis elegans." Physiol Rev **83**(2): 377-415.

Thiemann, A., S. Grunder, et al. (1992). "A chloride channel widely expressed in epithelial and non-epithelial cells." Nature **356**(6364): 57-60.

Townley, R. and L. Shapiro (2007). "Crystal structures of the adenylate sensor from fission yeast AMP-activated protein kinase." Science **315**(5819): 1726-1729.

Tseng, P. Y., B. Bennetts, et al. (2007). "Cytoplasmic ATP inhibition of CLC-1 is enhanced by low pH." J Gen Physiol **130**(2): 217-221.

von Rozycki, T., D. H. Nies, et al. (2005). "Genomic Analyses of Transport Proteins in Ralstonia metallidurans." Comp Funct Genomics **6**(1-2): 17-56.

Wang, D. N., M. Safferling, et al. (2003). "Practical aspects of overexpressing bacterial secondary membrane transporters for structural studies." Biochim Biophys Acta **1610**(1): 23-36.

Warmuth, S. (2009). "Biochemical and structural characterization of Cation Chloride Cotransporters " PhD thesis. University of Zurich.

White, M. M. and C. Miller (1979). "A voltage-gated anion channel from the electric organ of Torpedo californica." J Biol Chem **254**(20): 10161-10166.

Yamashita, A., S. K. Singh, et al. (2005). "Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters." Nature **437**(7056): 215-223.

Yang, Z. R., R. Thomson, et al. (2005). "RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins." Bioinformatics **21**(16): 3369-3376.

Zagotta, W. N., N. B. Olivier, et al. (2003). "Structural basis for modulation and agonist specificity of HCN pacemaker channels." Nature **425**(6954): 200-205.

Zhang, X. D., P. Y. Tseng, et al. (2008). "ATP inhibition of CLC-1 is controlled by oxidation and reduction." J Gen Physiol **132**(4): 421-428.

Zifarelli, G. and M. Pusch (2009). "Intracellular regulation of human ClC-5 by adenine nucleotides." EMBO Rep **10**(10): 1111-1116.

# List of abbreviations

| | |
|---|---|
| ADP | Adenosine Di-Phosphate |
| AMP | Adenosine Mono-Phosphate |
| ATP | Adenosine Tri-Phosphate |
| AU | Asymmetric Unit |
| AUC | Analytical Ultracentrifugation |
| CBS | Cystathionine β-Synthase |
| Da | Dalton |
| DDM | Dodecyl Maltoside |
| DM | Decyl Maltoside |
| DNA | Deoxyribonucleic Acid |
| DSMZ | Deutsche Sammlung von Mikroorganismen und Zellkulturen |
| FRET | Förster Resonance Energy Transfer |
| FX | Fragment Exchange |
| GFP | Green Fluorescent Protein |
| HGPRT | Hypoxanthine-Guanine Phosphoribosyltransferase |
| His | Histidine |
| HRP | Horseradish Peroxidase |
| IgG | Immunoglobulin G |
| IMAC | Immobilized Metal Affinity Chromatography |
| ITC | Isothermal Titration Calorimetry |
| LB | Lysogeny Broth |
| MWCO | Molecular Weight Cut Off |
| Ni-NTA | Nickel Nitrilotriacetic Acid |
| NM | Nonyl Maltoside |
| NMR | Nuclear Magnetic Resonance |
| OD | Optical Density |
| OM | Octyl Maltoside |
| PAGE | Polyacrylamid Gel Electrophoresis |
| PBS | Phosphate Buffered Saline |
| PCR | Polymerase Chain Reaction |
| PEG | Polyethylene Glycol |
| pS | Picosiemens |

| | |
|---|---|
| PVDF | Poly Vinyl Difluoride |
| SDS | Sodium Dodecyl Sulphate |
| SEC | Size Exclusion Chromatography |
| TB | Terrific Broth |
| TBS | Tris Buffered Saline |
| WT | Wild Type |

# Curriculum Vitae

**Sandra MARKOVIC**

| | |
|---|---|
| Date of birth | November 28th, 1978 |
| Place of birth | Jagodina, Serbia |
| Nationality | Serbian |

## Education

| | |
|---|---|
| 01/2006 – 04/2009 | Ph.D. thesis at the University of Zurich and Molecular Life Science PhD Program under the supervision of Prof. Raimund Dutzler |
| 09/2002-06/2005 | Master thesis at the Institute of Molecular Genetics and Genetic Engineering, Faculty of Biology, University of Belgrade, Serbia; Title: "Analysis of autoregulation of KgmB methylase from *Streptomyces tenebrarius*" |
| 10/2001- 03/2002 | Diploma thesis, Faculty of Chemistry, University of Belgrade; Title: "Production and characterization of pectinases from *Aspergillus niger*" |
| 1997-2002 | Biochemistry studies, Faculty of Chemistry, University of Belgrade, Serbia |
| 1993-1997 | Gymnasium Zrenjanin, Serbia |

## Professional Development

| | |
|---|---|
| 06/2009 | Poster presentation - VIII European Symposium of the Protein Society, Zurich, Switzerland; Title: "The structure of the cytoplasmic domain of the chloride channel ClC-Ka reveals a conserved interaction interface" |

| | |
|---|---|
| 09/2008 | The 9th International School on the Crystallography of Biological Macromolecules", Como, Italy |
| 07/2008 | Poster presentation - Ion Channels Gordon Research Conference, Tilton, USA; Title: "The structure of the cytoplasmic domain of the chloride channel ClC-Ka reveals a conserved interaction interface" |
| 09/2007 | Oral presentation - 21st Rhine-Knee Regiomeeting, Hölstein, Switzerland; Title: "The structure of the cytoplasmic domain of the chloride channel ClC-Ka reveals a conserved interaction interface" |
| 03/2007 | Postgraduate Course – Zurich Center for Integrative Human Physiology: Title: "Electrophysiology" |
| 2006-2009 | 5th, 6th and 7th International NCCR Symposia on New Trends in Structural Biology, Zurich, Switzerland |

## Publication

**Markovic, S.** and Dutzler, R. The structure of the cytoplasmic domain of the chloride channel ClC-Ka reveals a conserved interaction interface. Structure 15, 715-725, (2007)